PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-352994
(43)Date of publication of application : 24.12.1999

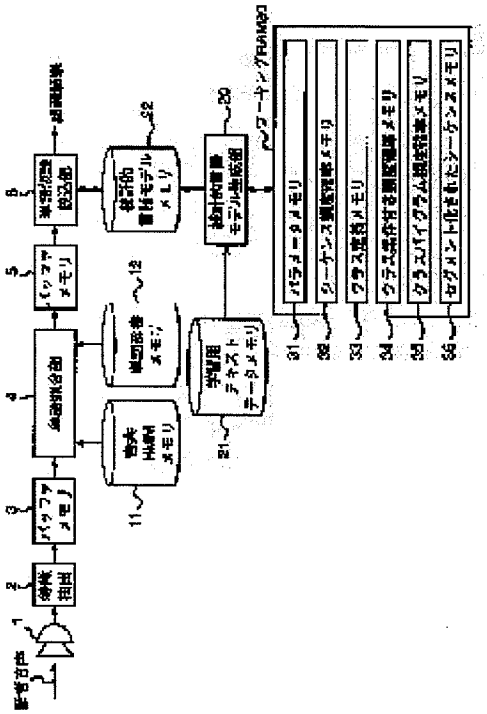
(51)Int.Cl. G10L 3/00
G10L 3/00
G10L 3/00
// C12N 15/09

(21)Application number : 10-165030
(22)Date of filing : 12.06.1998

(71)Applicant : ATR ONSEI HONYAKU TSUSHIN KENKYUSHO:KK
(72)Inventor : SABIN DERIN
KOSAKA YOSHINORI
NAKAJIMA HIDEJI

(54) STATISTICAL SEQUENCE MODEL GENERATOR, STATISTICAL LANGUAGE MODEL GENERATOR, AND SPEECH RECOGNITION SYSTEM

(57)Abstract:
PROBLEM TO BE SOLVED: To impart degree of freedom to an analyzed result, and to process a variable length of sequence as the same class, by re-estimating frequency probability of a bigram between sequences based on front likelihood, frequency probability, and rear likelihood.
SOLUTION: Re-estimation is conducted in a statistical language model generating part 20 to provide the maximum likelihood estimate using EM algorithm based on frequency probability for character lines included in sorted classes and conditional character lines and frequency probability for a bigram between the classes. Frequency probability of a bigram between sequences is re-estimated using an expression for expressing the frequency probability of the bigram between the sequences based on front likelihood for the character lines of processing objects put in a front side of a time series, frequency probability of the character lines when the character line just before the character lines is conditioned, and rear likelihood corresponding to the character lines put in a rear side of the time series, relating to the respective character lines of the processing objects. Bi-multigram statistical sequence models are generated to be output.



(19)日本国特許庁（ＪＰ）　(12)特　許　公　報（Ｂ２）　(11)特許番号
特許第3004254号
(P3004254)
(45)発行日　平成12年１月31日(2000.1.31)　(24)登録日　平成11年11月19日(1999.11.19)

(51)Int.Cl.[7]	識別記号	ＦＩ	
Ｇ１０Ｌ　15/28		Ｇ１０Ｌ　3/00	５７１Ｄ
15/06			５２１Ｃ
15/14			５３５Ｚ
// Ｃ１２Ｎ　15/09		Ｃ１２Ｎ　15/00	Ａ

請求項の数９(全 24 頁)

(21)出願番号　特願平10－165030
(22)出願日　平成10年６月12日(1998.6.12)
(65)公開番号　特開平11－352994
(43)公開日　平成11年12月24日(1999.12.24)
審査請求日　平成10年６月12日(1998.6.12)

(73)特許権者　593118597
株式会社エイ・ティ・アール音声翻訳通信研究所
京都府相楽郡精華町大字乾谷小字三平谷５番地
(72)発明者　サビン・デリン
京都府相楽郡精華町大字乾谷小字三平谷５番地　株式会社エイ・ティ・アール音声翻訳通信研究所内
(72)発明者　匂坂　芳典
京都府相楽郡精華町大字乾谷小字三平谷５番地　株式会社エイ・ティ・アール音声翻訳通信研究所内
(74)代理人　100062144
弁理士　青山　葆　(外２名)

審査官　涌井　智則

最終頁に続く

(54)【発明の名称】　統計的シーケンスモデル生成装置、統計的言語モデル生成装置及び音声認識装置

(57)【特許請求の範囲】
【請求項１】　１個又は複数の単位からなる単位列であるシーケンスを含む入力データに基づいて、可変長の自然数N_1個の単位列であるマルチグラムと、可変長の自然数N_2個の単位列であるマルチグラムとの間のバイグラムであるバイーマルチグラムの統計的シーケンスモデルを生成する統計的シーケンスモデル生成装置であって、
上記入力データに基づいて、予め決められたN_1，N_2の最大値の拘束条件のもとで、すべての単位列の組み合わせの上記バイグラムの頻度確率を計数する初期化手段と、
上記初期化手段によって計数された上記バイグラムの頻度確率に基づいて、各クラスの対をマージしたときの相互情報量の損失が最小となるようにマージして各クラス

の頻度確率を更新して予め決められた数の複数のクラスに分類することにより、分類されたクラスに含まれる単位列と、分類されたクラスの条件付きの単位列の頻度確率と、分類されたクラス間のバイグラムの頻度確率を計算して出力する分類手段と、
上記分類処理手段から出力される分類されたクラスに含まれる単位列と、分類されたクラスの条件付きの単位列の頻度確率と、分類されたクラス間のバイグラムの頻度確率とに基づいて、ＥＭアルゴリズムを用いて、最尤推定値を得るように再推定し、ここで、フォワード・バックワードアルゴリズムを用いて、処理対象の各単位列に対して、時系列的に前方にとり得る処理対象の当該単位列に対する前方尤度と、当該単位列の直前の単位列を条件としたときの当該単位列の頻度確率と、時系列的に後方にとり得る当該単位列に対する後方尤度とに基づいて

シーケンス間のバイグラムの頻度確率を示す式を用いて、当該シーケンス間のバイグラムの頻度確率を再推定することにより、再推定結果である上記バイ－マルチグラムの統計的シーケンスモデルを生成して出力する再推定手段と、

上記分類手段の処理と上記再推定手段の処理を所定の終了条件を満たすまで繰り返し実行するように制御する制御手段とを備えたことを特徴とする統計的シーケンスモデル生成装置。

【請求項２】　上記初期化手段はさらに、上記計数されたバイグラムの頻度確率のうち、所定の頻度確率以下のバイグラムの組み合わせのデータを除去することを特徴とする請求項１記載の統計的シーケンスモデル生成装置。

【請求項３】　上記分類手段は、上記初期化手段によって計数された上記バイグラムの頻度確率に基づいて、ブラウンアルゴリズムを用いて、上記複数のクラスに分類することを特徴とする請求項１又は２記載の統計的シーケンスモデル生成装置。

【請求項４】　上記式は、上記入力データにおいて、当該単位列である第２の単位列が第１の単位列に続くときの単位列のシーケンス間のバイグラムの頻度確率を、上記入力データにおける処理対象の各単位列に対して計算するための式であり、

上記シーケンス間のバイグラムの頻度確率は、第１と第２の単位列を含むすべてのセグメント化での尤度の和を、第１の単位列を含むすべてのセグメント化での尤度の和で除算することによって得られたことを特徴とする請求項１乃至３のうちの１つに記載の統計的シーケンスモデル生成装置。

【請求項５】　上記式は、上記入力データにおいて各単位列が発生する平均回数を示す分母と、上記入力データにおいて第２の単位列が第１の単位列に続くときの各単位列に対する平均回数を示す分子とを有し、

上記分子は、処理対象の各単位列に対する、上記前方尤度と、当該単位列の直前の単位列を条件としたときの当該単位列の頻度確率と、上記後方尤度の積の和であり、

上記分母は、処理対象の各単位列に対する、上記前方尤度と、当該単位列の直前の単位列を条件としたときのすべての単位列の頻度確率と、上記後方尤度の積の和であることを特徴とする請求項４記載の統計的シーケンスモデル生成装置。

【請求項６】　上記終了条件は、上記分類手段の処理と、上記再推定手段の処理との反復回数が予め決められた回数に達したときであることを特徴とする請求項１乃至５のうちの１つに記載の統計的シーケンスモデル生成装置。

【請求項７】　請求項１乃至６のうちの１つに記載の統計的シーケンスモデル生成装置において、

上記単位は自然言語の文字であり、上記シーケンスは単語であり、上記分類手段は、文字列を複数の単語の列に分類し、上記統計的シーケンスモデルは、統計的言語モデルであることを特徴とする統計的言語モデル生成装置。

【請求項８】　請求項１乃至６のうちの１つに記載の統計的シーケンスモデル生成装置において、上記単位は自然言語の単語であり、上記シーケンスはフレーズであり、上記分類手段は、単語列を複数のフレーズの列に分類し、上記統計的シーケンスモデルは、統計的言語モデルであることを特徴とする統計的言語モデル生成装置。

【請求項９】　入力される発声音声文の音声信号に基づいて、所定の統計的言語モデルを用いて音声認識する音声認識手段を備えた音声認識装置において、上記音声認識手段は、請求項７又は８記載の統計的言語モデル生成装置によって生成された統計的言語モデルを参照して音声認識することを特徴とする音声認識装置。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、学習用シーケンスデータに基づいて統計的シーケンスモデルを生成する統計的シーケンスモデル生成装置、学習用テキストデータに基づいて統計的言語モデルを生成する統計的言語モデル生成装置、及び上記統計的言語モデルを用いて、入力される発声音声文の音声信号を音声認識する音声認識装置に関する。

【０００２】

【従来の技術】近年、連続音声認識装置において、その性能を高めるために言語モデルを用いる方法が研究されている。これは、シーケンスモデルである言語モデルを用いて、次単語を予測し探索空間を削減することにより、認識率の向上及び計算時間の削減の効果を狙ったものである。ここで、シーケンスとは、具体的には、文字のシーケンスでは単語であり、単語のシーケンスではフレーズ（又は句）である。最近盛んに用いられている言語モデルとしてＮ－ｇｒａｍ（Ｎ－グラム；ここで、Ｎは２以上の自然数である。）がある。これは、大規模なテキストデータを学習し、直前のＮ－１個の単語から次の単語への遷移確率を統計的に与えるものである。複数Ｌ個の単語列$w_1^L = w_1, w_2, \cdots, w_L$の生成確率Ｐ（$w_1^L$）は次式で表される。

【０００３】

【数１】

$$P(w_1^L) = \prod_{t=1}^{L} P(w_t \mid w_{t+1-N}^{t-1})$$

【０００４】ここで、w_tは単語列w_1^Lのうちｔ番目の１つの単語を表し、w_i^jはｉ番目からｊ番目の単語列を表わす。上記数１において、確率Ｐ（$w_t \mid w_{t-1-N}^{t-1}$）は、Ｎ個の単語からなる単語列w_{t-1-N}^{t-1}

が発声された後に単語w_tが発声される確率であり、以下同様に、確率Ｐ（Ａ｜Ｂ）は単語又は単語列Ｂが発声された後に単語Ａが発声される確率を意味する。また、数１における「Π」はｔ＝１からＬまでの確率Ｐ（w_t｜w_{t+1-N}^{t-1}）の積を意味する。

【０００５】ところで、近年、上記統計的言語モデルのＮ－ｇｒａｍを用いて連続音声認識の性能を向上させる手法が盛んに提案されており、そのいくつかのモデルでは、可変長の単語列にわたる単語の依存性を利用する方法を用いている。これらのモデルは、共通して従来のＮ－ｇｒａｍモデルにみられる固定長の依存性の仮定を緩和するために用いられており、種々のより広い仮定をカバーしている。

【０００６】フレーズを純粋に統計的方法（すなわち、統計的文脈自由文法（Stochastic Context Free Grammars）にあるような文法的規則を用いない方法）で導くためには、種々の基準を使用する必要があり、例えば、以下の基準が提案されてきた。

（ａ）従来技術文献１「K. Ries et al.,”Class phrase models for languagemodeling”, Proceedings of ICSLP 96, 1996」において開示されたリーブ・ワン・アウト（leave-one-out）尤度、及び

（ｂ）従来技術文献２「H. Masataki et al., Variable-order n-gram generation by word-class splitting and consecutive word grouping. Proceedings ofICASSP 96,1996」において開示されたエントロピー。

【０００７】

【発明が解決しようとする課題】これらの方法において、尤度の基準を統計的枠組みの中で用いることで、ＥＭ（Expectation Maximum;すなわち、期待値の最大化）アルゴリズムを用いた最適化の方法を用いることができるが、過学習となる傾向がある。また、最適化処理においては、例えば、従来技術文献３「S. Matsunaga et al.,”Variable-length language modeling integrating global constraints”,Proceedings of EUROSPEECH 97,1997」において発見的手法を用いられているが、統計的言語モデルの収束と最適化は理論的に保証されていない。

【０００８】ここで、さらに、例えば、従来技術文献１において提案された尤度の基準を用いたときの問題点について述べると以下の通りである。

＜問題点１＞単語のシーケンスの頻度確率が貪欲なアルゴリズム(greedy algorithm)によって得られるために、最適な状態に向かう単調な収束が保証されない。

＜問題点２＞この方法は確定的なものである。つまり、仮にシーケンス［ｂｃｄ］がシーケンスの目録（inventory)に在れば、入力文字列に”ｂｃｄ”が発生しても、これが［ｂｃ］＋［ｄ］、［ｂ］＋［ｃｄ］、［ｂ］＋［ｃ］＋［ｄ］等のサブシーケンスに分割されることはない。言い換えれば、シーケンスへの解析において自由度が無い。

＜問題点３＞シーケンスのクラスの定義が先行する単語のクラス分類を基礎としている。すなわち、まず、単語が分類され、次に、単語のクラスのラベルの各シーケンスは、シーケンスのクラスを定義するために使用される。従って、同一クラスに長さの違うシーケンスを入れることはできない。例えば、”thank you for”と”thank you very much for”は同じクラスに入らない。

【０００９】これを解決するために、本発明者は、従来技術文献４「S. Deligne et al.,”Introducing statistical dependencies and structural constraints in variable-length sequence models”、In Grammatical Inference: Learning Syntaxfrom Sentences, Lecture Notes in Artificial Intelligence 1147, pp.156-167,Springer,1996」において、可変長のシーケンスであるマルチグラムを用いる統計的言語モデルについて、当該従来技術文献４の（１６）式を用いて、それらのパラメータを計算できる可能性だけを示しているが、当該（１６）式は、実際にディジタル計算機を用いて計算することができる形式とはなっておらず、実用化することができないという問題点があった。ここで、マルチグラムとは、他のシーケンスとの依存性を特定しない可変長のシーケンスである。

【００１０】本発明の目的は以上の問題点を解決し、従来例に比較して、最適な状態に向かう単調な収束を保証することができ、解析結果に自由度があり、可変長のシーケンスを同一のクラスで取り扱うことができ、ディジタル計算機を用いて実用的に高速処理して統計的モデルを生成することができる統計的シーケンスモデル生成装置、統計的言語モデル生成装置及び音声認識装置を提供することにある。

【００１１】

【課題を解決するための手段】本発明に係る統計的シーケンスモデル生成装置は、１個又は複数の単位からなる単位列であるシーケンスを含む入力データに基づいて、可変長の自然数N_1個の単位列であるマルチグラムと、可変長の自然数N_2個の単位列であるマルチグラムとの間のバイグラムであるバイーマルチグラムの統計的シーケンスモデルを生成する統計的シーケンスモデル生成装置であって、上記入力データに基づいて、予め決められたN_1，N_2の最大値の拘束条件のもとで、すべての単位列の組み合わせの上記バイグラムの頻度確率を計数する初期化手段と、上記初期化手段によって計数された上記バイグラムの頻度確率に基づいて、各クラスの対をマージしたときの相互情報量の損失が最小となるようにマージして各クラスの頻度確率を更新して予め決められた数の複数のクラスに分類することにより、分類されたクラスに含まれる単位列と、分類されたクラスの条件付きの単位列の頻度確率と、分類されたクラス間のバイグラムの頻度確率を計算して出力する分類手段と、上記分類処

理手段から出力される分類されたクラスに含まれる単位列と、分類されたクラスの条件付きの単位列の頻度確率と、分類されたクラス間のバイグラムの頻度確率とに基づいて、ＥＭアルゴリズムを用いて、最尤推定値を得るように再推定し、ここで、フォワード・バックワードアルゴリズムを用いて、処理対象の各単位列に対して、時系列的に前方にとり得る処理対象の当該単位列に対する前方尤度と、当該単位列の直前の単位列を条件としたときの当該単位列の頻度確率と、時系列的に後方にとり得る当該単位列に対する後方尤度とに基づいてシーケンス間のバイグラムの頻度確率を示す式を用いて、当該シーケンス間のバイグラムの頻度確率を再推定することにより、再推定結果である上記バイーマルチグラムの統計的シーケンスモデルを生成して出力する再推定手段と、上記分類手段の処理と上記再推定手段の処理を所定の終了条件を満たすまで繰り返し実行するように制御する制御手段とを備えたことを特徴とする。

【００１２】また、上記統計的シーケンスモデル生成装置において、上記初期化手段はさらに、上記計数されたバイグラムの頻度確率のうち、所定の頻度確率以下のバイグラムの組み合わせのデータを除去することを特徴とする。

【００１３】さらに、上記統計的シーケンスモデル生成装置において、上記分類手段は、上記初期化手段によって計数された上記バイグラムの頻度確率に基づいて、ブラウンアルゴリズムを用いて、上記複数のクラスに分類することを特徴とする。

【００１４】また、上記統計的シーケンスモデル生成装置において、上記式は、上記入力データにおいて、当該単位列である第２の単位列が第１の単位列に続くときの単位列のシーケンス間のバイグラムの頻度確率を、上記入力データにおける処理対象の各単位列に対して計算するための式であり、上記シーケンス間のバイグラムの頻度確率は、第１と第２の単位列を含むすべてのセグメント化での尤度の和を、第１の単位列を含むすべてのセグメント化での尤度の和で除算することによって得られる。また、ここで、上記式は、上記入力データにおいて各単位列が発生する平均回数を示す分母と、上記入力データにおいて第２の単位列が第１の単位列に続くときの各単位列に対する平均回数を示す分子とを有し、上記分子は、処理対象の各単位列に対する、上記前方尤度と、当該単位列の直前の単位列を条件としたときの当該単位列の頻度確率と、上記後方尤度の積の和であり、上記分母は、処理対象の各単位列に対する、上記前方尤度と、当該単位列の直前の単位列を条件としたときのすべての単位列の頻度確率と、上記後方尤度の積の和である。

【００１５】さらに、上記統計的シーケンスモデル生成装置において、上記終了条件は、上記分類手段の処理と、上記再推定手段の処理との反復回数が予め決められた回数に達したときであることを特徴とする。

【００１６】また、本発明に係る統計的言語モデル生成装置は、上記統計的シーケンスモデル生成装置において、上記単位は自然言語の文字であり、上記シーケンスは単語であり、上記分類手段は、文字列を複数の単語の列に分類し、上記統計的シーケンスモデルは、統計的言語モデルであることを特徴とする。

【００１７】さらに、本発明に係る統計的言語モデル生成装置は、上記統計的シーケンスモデル生成装置において、上記単位は自然言語の単語であり、上記シーケンスはフレーズであり、上記分類手段は、単語列を複数のフレーズの列に分類し、上記統計的シーケンスモデルは、統計的言語モデルであることを特徴とする。

【００１８】またさらに、本発明に係る音声認識装置は、入力される発声音声文の音声信号に基づいて、所定の統計的言語モデルを用いて音声認識する音声認識手段を備えた音声認識装置において、上記音声認識手段は、上記統計的言語モデル生成装置によって生成された統計的言語モデルを参照して音声認識することを特徴とする。

【００１９】

【発明の実施の形態】以下、図面を参照して本発明に係る実施形態について説明する。以下の実施形態においては、単位は文字であり、文字のシーケンスである文字列を単語列に分類する一例、並びに、単位は単語であり、単語のシーケンスである単語列をフレーズ（句）に分類する一例について説明しているが、本発明はこれに限らず、単位はＤＮＡであり、ＤＮＡのシーケンスであるＤＮＡ列を所定のＤＮＡ配列に分類するように構成してもよい。また、単位は塩基であり、塩基のシーケンスである塩基列を所定のコドンに分類するように構成してもよい。

【００２０】図１は、本発明に係る一実施形態である連続音声認識装置のブロック図である。本実施形態の連続音声認識装置は、学習用テキストデータメモリ２１に記憶された文字列であるテキストデータに基づいて、ワーキングＲＡＭ３０を用いて、可変長のバイーマルチグラムの言語モデルを生成する統計的言語モデル生成部２０を備え、ここで、統計的言語モデル生成部２０の処理は、図３に示すように、大きく分けると、ブラウンアルゴリズムを用いた分類処理（ステップＳ３）と、バイーマルチグラムを用いた再推定処理（ステップＳ４）とを含むことを特徴としている。

【００２１】すなわち、本実施形態の統計的言語モデル生成装置は、１個又は複数の文字からなる文字列のシーケンスを含む入力データに基づいて、可変長の自然数N_1個の文字列と可変長の自然数N_2個の文字列との間のバイグラムであるバイーマルチグラムの統計的言語モデルを生成する統計的言語モデル生成装置であり、ここで、図３に示すように、（ａ）上記入力データに基づいて、予め決められたN_1，N_2の最大値の拘束条件のもとで、

すべての文字列の組み合わせの上記バイグラムの頻度確率を計数する初期化処理（ステップＳ２）と、（ｂ）上記初期化処理によって計数された上記バイグラムの頻度確率に基づいて、各クラスの対をマージしたときの相互情報量の損失が最小となるようにマージして各クラスの頻度確率を更新して予め決められた数の複数のクラスに分類することにより、分類されたクラスに含まれる文字列と、分類されたクラスの条件付きの文字列の頻度確率と、分類されたクラス間のバイグラムの頻度確率を計算して出力する分類処理（ステップＳ３）と、（ｃ）上記分類処理によって得られた分類されたクラスに含まれる文字列と、分類されたクラスの条件付きの文字列の頻度確率と、分類されたクラス間のバイグラムの頻度確率とに基づいて、ＥＭアルゴリズムを用いて、最尤推定値を得るように再推定し、ここで、フォワード・バックワードアルゴリズムを用いて、処理対象の各文字列に対して、時系列的に前方にとり得る処理対象の当該文字列に対する前方尤度と、当該文字列の直前の文字列を条件としたときの当該文字列の頻度確率と、時系列的に後方にとり得る当該文字列に対する後方尤度とに基づいてシーケンス間のバイグラムの頻度確率を示す式（数２２－数２４）を用いて、当該シーケンス間のバイグラムの頻度確率を再推定することにより、再推定結果である上記バイーマルチグラムの統計的シーケンスモデルを生成して出力する再推定処理（ステップＳ４）と、（ｄ）上記分類処理と上記再推定処理を所定の終了条件を満たすまで繰り返し実行するように制御する処理（ステップＳ５）を含むことを特徴とする。

【００２２】本実施形態では、単語のＮ－ｇｒａｍに基づく手法に対向する、フレーズに基づく方法に焦点を当てる。ここで、複数の文はフレーズに構成され、頻度確率は、単語に代わってフレーズに割り当てられる。モデルがＮ－ｇｒａｍに基づくか、フレーズに基づくかに関わらず、それらは確定的モデルあるいは統計的モデルのいずれかに該当する。フレーズに基づく枠組みでは、非確定性はその文の解析結果の曖昧さを通じてフレーズに導入される。すなわち、これは実際においては、フレーズ”ａｂｃ”がフレーズとして登録されているにもかかわらず、文字列の解析結果が例えば［ａｂ］［ｃ］となる確率が皆無でないことを意味する。これとは対照的に、確定的手法ではａ、ｂ、ｃすべての同時出現はシステマティックにフレーズ［ａｂｃ］の出現と解釈される。

【００２３】また、本実施形態では、統計的言語モデルの処理は、バイーマルチグラムを用いて実行され、当該バイーマルチグラムの言語モデルは、フレーズに基づく統計的モデルであり、そのパラメータは尤度基準に従って推定される。

【００２４】まず、マルチグラムの理論的な定式化につ＊いて説明する。マルチグラムの枠組みでは、Ｔ個の単語からなる文

【数２】$W=w_{(1)}\ w_{(2)}\ \cdots w_{(T)}$

は、それぞれ最大長ｎ個の単語からなる各々のフレーズが連鎖（シーケンス）したものと仮定される。ここで、ＳはＴs個のフレーズへのセグメント化を示し、$s_{(t)}$ はセグメント化Ｓにおける時刻インデックス（最初の語からのシリアル番号を示す。）（ｔ）のフレーズとした場合、Ｗのセグメント化の結果は、次式で表すことができる。

【数３】$(W,\ S)=s_{(1)}\ \cdots s_{(Ts)}$

【００２５】ここで、セグメント化された複数のフレーズからなる辞書は、語彙から１，２…からｎにいたるまでの単語を組み合わせて形成されるものであり、ここでは，次式のように表す。

【数４】$Ds=\{s_j\}_j$

そして、文の尤度は、各セグメント化に対する尤度の和として、次式のように計算される。

【００２６】

【数５】

$$L(W)=\sum_{S\in\{S\}} L(W,S)$$

【００２７】モデルの決定指向的手法により、文Ｗは、最も尤らしいセグメント化に従って解析され、次の近似式が得られる。

【００２８】

【数６】

$$L^{*}(W)=\max_{S\in\{S\}} L(W,S)$$

【００２９】ここで、フレーズ間のｎ－ｇｒａｍの相関を仮定し、特定のセグメント化Ｓの結果の尤度の値を次式のように計算する。

【００３０】

【数７】

$$L(W,S)=\prod_{t} p\left(s_{(t)} \mid s_{(t-n+1)}\cdots s_{(t-1)}\right)$$

【００３１】ここで、以下、符号ｎは複数のフレーズ間の依存度を表し、従来のｎ－ｇｒａｍの表記法のｎとして使用する。また、符号n_{max} は、フレーズの最大長を表す。従って、ここで、尤度の計算例を次式に示す。この例では、バイーマルチグラムモデル（$n_{max}=3$，$n=2$）の”ａｂｃｄ”の尤度を示す。記号＃は空のシーケンスを表す。

【００３２】

【数８】

尤度

11　　12

＝ｐ（［ａ］｜＃）ｐ（［ｂ］｜［ａ］）ｐ（［ｃ］｜［ｂ］）ｐ（［ｄ］｜［ｃ］）
＋ｐ（［ａ］｜＃）ｐ（［ｂ］｜［ａ］）ｐ（［ｃｄ］｜［ｂ］）
＋ｐ（［ａ］｜＃）ｐ（［ｂｃ］｜［ａ］）ｐ（［ｄ］｜［ｂｃ］）
＋ｐ（［ａ］｜＃）ｐ（［ｂｃｄ］｜［ａ］）
＋ｐ（［ａｂ］｜＃）ｐ（［ｃ］｜［ａｂ］）ｐ（［ｄ］｜［ｃ］）
＋ｐ（［ａｂ］｜＃）ｐ（［ｃｄ］｜［ａｂ］）
＋ｐ（［ａｂｃ］｜＃）ｐ（［ｄ］｜［ａｂｃ］）

【００３３】上記数８から明らかなように、当該尤度は、シーケンス”ａｂｃｄ”をセグメント化するときのすべての組み合わせについての頻度確率の和を表している。

【００３４】次いで、言語モデルのパラメータの推定について説明する。マルチグラムのｎ－ｇｒａｍモデルは、パラメータΘのセットによって完全に定義され、次式のパラメータΘは、辞書Ｄｓを用いて、

【数９】

$$\Theta = \{ p(s_{in} \mid s_{i1} \cdots s_{in-1}) \mid s_{i1} \cdots s_{in} \in Ds \}$$

ｎ個のフレーズのあらゆる組み合わせに関係するｎ－ｇｒａｍの条件付き確率によって構成される。パラメータ＊Θのセットの推定値は、例えば、不完全なデータから得られる想定しうる最大の尤度値、すなわち最尤推定値（Maximum Likelihood Estimation）として得られ、ここで、未知のデータは基礎をなすセグメント化Ｓである。従って、パラメータΘの反復的な最尤推定値は、公知のＥＭアルゴリズム（Expectation Maximization Algorithm)によって計算することができる。ここで、Ｑ（ｋ，ｋ＋１）を、反復回数パラメータｋ及びｋ＋１の尤度を用いて計算される、次式の補助関数とする。

【００３５】

【数１０】

$$Q(k, k+1) = \sum_{S \in \{S\}} L^{(k)}(S \mid W) \log \{ L^{(k+1)}(W, S) \}$$

【００３６】公知のＥＭアルゴリズムにおいて示されるように、

【数１１】Ｑ（ｋ，ｋ＋１）≧Ｑ（ｋ，ｋ）

であれば、

【数１２】$L^{(k+1)}(W) \geqq L^{(k)}(W)$

である。従って、反復回数パラメータ（ｋ＋１）における次式の再推定式

【数１３】$p^{(k+1)}(s_{in} \mid s_{i1} \cdots s_{in-1})$

は、次式の拘束条件

【数１４】※

$$\sum_{s_{ib} \in D_s} p(s_{ib} \mid s_{i1} \cdots s_{ib-1}) = 1$$

のもとで、モデルパラメータ$\Theta^{(k+1)}$について補助関数Ｑ（ｋ，ｋ＋１）を最大化することにより、次式のように直接的に導くことができる。なお、本明細書において、下付きの下付きの表記及び上付きの下付きの表記はできないので、下層の下付きの表記を省略している。

【００３７】

【数１５】※

$$p^{(k+1)}(s_{ib} \mid s_{i1} \cdots s_{ib-1}) = p_a / p_b$$

ここで、

$$p_a = \{ \sum_{S \in \{S\}} c(s_{i1} \cdots s_{ib-1} s_{ib}, S) \times L^{(k)}(S \mid W) \}$$

$$p_b = \{ \sum_{S \in \{S\}} c(s_{i1} \cdots s_{ib-1}, S) \times L^{(k)}(S \mid W) \}$$

【００３８】ここで、ｃ（$s_{i1} \cdots s_{in}$，Ｓ）は、セグメント化Ｓにおける複数のフレーズ$s_{i1} \cdots s_{in}$の組み合わせの出現数を示す。数１５の再推定式は、バイーマルチグラム（ｎ＝２）について詳細後述されるように、フォワード・バックワードアルゴリズム（forward backward algorithm)(以下、ＦＢ法ともいう。）を用いて実行される。決定指向の方法では、再推定式は、次式のように簡略化される。

【００３９】

【数１６】$p^{(k-1)}(s_{in} \cdots s_{in-1}) = \{ c(s_{i1} \cdots s_{in-1} s_{in}, S^{*(k)}) \} / \{ c(s_{i1} \cdots s_{in-1}, S^{*(k)}) \}$

【００４０】ここで、$S^{*(k)}$は、$L^{(k)}(S \mid W)$を最大化する文の解析結果であり、ビタビ（Viterbi）アルゴリズムによって導かれる。各反復は、尤度$L^{(k)}(W)$を増大させる意味において言語モデルを改善し、最終的には臨界点（おそらくは、局所最大値）へ収束する。モデルパラメータΘのセットは、学習用コーパス、すなわち学習用テキストデータにおいて観察されるすべてのフレーズの組み合わせの相対的頻度を用いて初期化され

る。

【００４１】次いで、可変長フレーズのクラスタリング（分類処理）について説明する。従来技術文献１によれば、近年、クラスーフレーズに基づくモデルが注目されているが、通常、それは従来の単語クラスタリングを仮定している。典型的には、各単語はまず、単語が属するクラスのラベルC_kを割り当てられ、単語ークラスラベルの可変長フレーズ［C_{k1}，C_{k2}…C_{kn}］が導かれる。各可変長フレーズによって、“＜［C_{k1}，C_{k2}…C_{kn}］＞”として示されるフレーズが属するクラスのラベルが定義される。しかしながら、この手法では、同じ長さのフレーズのみにしか同じフレーズークラスラベルを割り当てることができない。例えば、”thank you for”と”thank you very much for”というフレーズを同じクラスラベルに割り当てることができない。本実施形態では、このような限界に対する解決法として、単語に代わり直接フレーズをクラスタリングする方法を提案する。この目的を達成するためには、２個のフレーズ間のバイグラムの相関（$n_{max}=2$）を仮定し、上述したバイーマルチグラムモデルの学習手法に変更を加え、各反復が次の２つの段階より構成されるようにする。

【００４２】（Ｉ）ステップＳＳ１：クラス割り当て（図３のステップＳ３に対応する。）

＊【数１７】$\{p^{(k)}(s_j \mid s_i)\} \rightarrow \{p^{(k)}(C_{k(sj)} \mid C_{k(sj)}),\ p^{(k)}(s_j \mid C_{k(sj)})\}$

（ＩＩ）ステップＳＳ２：マルチグラムの再推定（図３のステップＳ４に対応する。）

【数１８】$\{p^{(k)}(C_{k(sj)} \mid C_{k(si)}),\ p^{(k)}(s_j \mid C_{k(sj)})\} \rightarrow \{p^{(k+1)}(s_j \mid s_i)\}$

【００４３】上記ステップＳＳ１では、フレーズバイグラムの頻度確率を入力とし、クラスバイグラムの頻度確率を出力する。クラス割り当ては、例えば、従来技術文献５「P. F. Brown et al., ”Class-based n-gram models of natural language”, Computational Linguistics, Vol. 18, No. 4, pp. 467-479, 1992」によれば、隣り合うフレーズ間の相関情報を最大化することによって行われる。ここで、クラスタリングの候補は単語ではなくフレーズとする。上述のように、$\{p^{(0)}(s_j \mid s_i)\}$は、学習用テキストデータにおけるフレーズの同時出現の相対的頻度を用いて初期化される。上記ステップＳＳ２では、マルチグラムの再推定式（数１５）又はその近似式（数１６）を用いてフレーズの頻度確率を再推定する。ここで、唯一の違いは、解析結果の尤度は以下の式により計算される。

【００４４】

＊【数１９】

$$L(W, S) = \prod_t p(C_{k(s(t))} \mid C_{k(s(t-1))})\, p(s_{(t)} \mid C_{k(s(t))})$$

【００４５】これは、上述したように、頻度確率$p^{(k)}(s_j \mid s_i)$に対する処理と同様に、頻度確率$p^{(k)}(C_{k(sj)} \mid C_{k(si)}) \times p^{(k)}(s_j \mid C_{k(sj)})$に基づいて頻度確率$p^{(k+1)}(s_j \mid s_i)$を再推定することに等しい。

【００４６】要約すれば、上記ステップＳＳ１によって、現在のフレーズ分布に関し、相互情報量の基準に基づくクラス割り当てが最適化されるよう保証され、上記ステップＳＳ２によって、現在のクラスの頻度確率を用いて、上記数１９に従って、計算された尤度がフレーズの頻度確率により最適化されるよう保証される。学習データは、従って、完全に統合化された方法により連合的（paradigmatic）かつ統合的（syntagmatic）（それぞれ言語学の用語である。）レベルの双方において反復的に構成される。すなわち、クラス割り当てにより表現されるフレーズ間の連合的関係はフレーズの頻度確率の再推定に影響を与え、フレーズの頻度確率は後続するクラス割り当てを決定する。

【００４７】本実施形態では、上述のように、バイーマルチグラムのパラメータの推定のために、フォワード・バックワードアルゴリズム（ＦＢ法）を用いる。これについて、以下に、詳述する。

【００４８】上記数１５は、フォワード・バックワードアルゴリズムを用いて、n_{max}をシーケンスの最大長とし、Ｔをコーパス（学習用テキストデータ）の語数として、複雑さの度合いであるコンプレキシティＯ（n_{max}^2Ｔ）で計算することができる。ここで、コンプレキシティＯ（n_{max}^2Ｔ）は計算コストのオーダーに対応する。すなわち、当該数１５の計算コストは、シーケンスの最大長n_{max}の２乗に比例し、コーパスの語数に比例する。本実施形態においては、基本的には、セグメント化｛Ｓ｝のセットではなく、単語のタイムインデックス（ｔ）にわたって加算を行い、数１５の分子及び分母を計算する。ここで、当該計算は、次式の前方向の変数$\alpha(t, l_i)$及び後ろ方向の変数$\beta(t, l_j)$の定義に依存する。

【００４９】

【数２０】

$\alpha(t, l_i) = L(W_{(t)}^{(t-1i)} \mid [W_{(t-1i+1)}^{(t)}])$

【数２１】

$\beta(t, l_j) = L(W_{(t+1)}^{(T)} \mid [W_{(t-1j-1)}^{(t)}])$

【００５０】前方向の変数$\alpha(t, l_i)$は、最初のｔ個の単語の尤度を表し、ここで、最後のl_i個の単語は、１つのシーケンスを形成するように制限される。また、後ろ方向の変数$\beta(t, l_j)$は、最後の（Ｔ－ｔ）個

15　　16

の語の条件付き尤度を示し、最後の（Ｔ－ｔ）個の単語は、シーケンス［$w_{(t-l_j+1)}$ …$w_{(t)}$ ］に後続する。ここで、例えば、$W_{(1)}^{(t-l_i)}$ は、時刻インデックス（１）から（ｔ－l_i）までの単語からなる単語列を表す。そして、解析結果の尤度は、数７によって計算されると仮定すると、数１５は次のように書き換えられる。

$$p_c=\sum_{t=1}^{T}\alpha(t,l_i)p^{(t)}(s_j\mid s_i)\beta(t+l_j,l_j)\delta_i(t-l_i+1)\delta_j(t+1)$$

【数２４】

$$p_d=\sum_{t}\alpha(t,l_i)\beta(t,l_i)\delta_i(t-l_i+1)$$

【００５２】ここで、l_i及びl_jはそれぞれシーケンスs_i及びs_jの長さを示す。クロネッカー関数δ_k（ｔ）は、時刻インデックスｔで開始する単語のシーケンスがs_kであるときは１となる一方、そうでない場合は０となる関数である。また、変数αおよびβは以下の反復式※（又は帰納式）によって計算できる。ここで、時刻インデックスｔ＝０及びｔ＝Ｔ＋１においてそれぞれ開始及び終了シンボルを仮定する。

【００５３】１≦ｔ≦Ｔ＋１に対して：

【数２５】

$$\alpha(t,l_i)=\sum_{l=1}^{n_{max}}\alpha(t-l_i,1)\,p\left([W_{(t-l_i+1)}^{(t)}]\mid[W_{(t-l_i-l+1)}^{(t-l_i)}]\right)$$

ここで、

【数２６】$\alpha(0,1)=1,\ \alpha(0,2)=\cdots=\alpha(0,n_{max})=0$ ★

である。

【００５４】０≦ｔ≦Ｔに対して：

【数２７】

$$\beta(t,l_i)=\sum_{l=1}^{n_{max}}p\left([W_{(t+1)}^{(t+1)}]\mid[W_{(t-l_i+1)}^{(t)}]\right)\beta(t+1,1)$$

ここで、

【数２８】$\beta(T+1,1)=1,\ \beta(T+1,2)=\cdots=\beta(T+1,n_{max})=0$

である。

【００５５】解析結果の尤度がクラスの仮定を用いて計算される場合、すなわち、数１９に従って計算される場合は、再推定式（数２２－数２４）の項$p^{(k)}(s_j\mid s_i)$はそのクラスの等価物、すなわち$p^{(k)}(C_{k(sj)}\mid C_{k(si)})\,p^{(k)}(s_j\mid C_{k(sj)})$に置き換えられる。αの反復式において、項$p([W_{(t-l_i+1)}^{(t)}]\mid[W_{(t-l_i-l+1)}^{(t-l_i)}])$は、シーケンス$[W_{(t-l_i-1)}^{(t)}]$のクラスの条件付き確率を乗算した対応するクラスのバイグラム確率に置き換えられる。同様の変形を反復式における変数βについても行う。

【００５６】次いで、本実施形態におけるフォワード・バックワードアルゴリズムを用いた再推定処理について、一例を参照して、以下に詳述する。前方向及び後ろ方向（以下、前後方向という。）の再推定処理は、数２２の分子の加算、及び分母の加算が、可能な解析結果集合｛Ｓ｝に代わって、学習データにおける単位の時刻インデックスｔについて計算されるように、数１５における複数の項を配列し直して行う。この方法は、前方向の変数α及び後ろ方向の変数βの定義に依存している。

（ａ）下記のパラグラフ＜＜Ａ１＞＞では、クラスのないことを仮定している。

（ｂ）下記のパラグラフ＜＜Ａ１．１＞＞では、変数α及びβを定義し、例を提供する。

（ｃ）下記のパラグラフ＜＜Ａ１．２＞＞では、変数α及びβを使用した頻度確率に関する前後方向の再推定について例示する。

（ｄ）下記のパラグラフ＜＜Ａ１．３＞＞では、反復（又は帰納）による変数αとβの計算方法に関して例示する。

（ｅ）下記のパラグラフ＜＜Ａ２＞＞では、クラスが存

在する場合のパラグラフ<<Ａ１．２>>及び<<Ａ１．３>>の修正方法を示す。

（ｆ）下記の例はすべて、次の表に示すデータに基づいて＊＊ている。

【００５７】

【表１】

入力学習データ（下記）：																								
o	n	e	s	i	x	o	n	e	e	i	g	h	t	s	i	x	t	h	r	e	e	t	w	o
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
単位の時刻インデックス（上記）：																								

（注）学習データの１つの文字は、１つの時刻インデックスに対応している。

【００５８】<<Ａ１．１>>前方向の変数αおよび後ろ方向の変数βの定義

変数α（ｔ，ｌ）は、長さｌのシーケンスで終了する、時刻インデックス（ｔ）までのデータの尤度である。例えば、変数α（９，３）は、シーケンス「ｏｎｅｓｉｘｏ_ｎ_ｅ」の尤度である。また、変数β（ｔ，ｌ）は、長さｌのシーケンスが時刻インデックス（ｔ）で終了するということが知られているときに、時刻インデックス（ｔ＋１）で開始されるデータの条件つき尤度である。例えば、変数β（９，３）は、先行するシーケンスが「ｏ_ｎ_ｅ」であるときの、シーケンス「ｅｉｇｈｔｓｉｘｔｈｒｅｅｔｗｏ」の尤度である。反復又は帰納による変数αおよびβの計算方法に関する例を、下記のパラグラフ<<Ａ１．３>>に示す。

【００５９】<<Ａ１．２>>変数αおよびβに基づく確率の再推定

例として、上記の学習データ例に関する、変数αおよびβを使用した頻度確率ｐ（ｏ_ｎ_ｅ｜ｓ_ｉ_ｘ）の再推定式を示す。頻度確率ｐ（ｏ_ｎ_ｅ｜ｓ_ｉ_ｘ）の一般的な再推定式（数１５））は次のような意味を持つ。

（ａ）分子は、学習データにおいてシーケンス「ｏ_ｎ_ｅ」がシーケンス「ｓ_ｉ_ｘ」に続く平均回数である。

（ｂ）分母は、学習データにおいてシーケンス「ｓ_ｉ_ｘ」が発生する平均回数である。

（ｃ）ここで、平均回数の値は、学習データのシーケンスにおけるすべての可能な解析結果について求める。

【００６０】フォワード・バックワードアルゴリズムを用いた再推定式（数２２－２４）の分子（数２３）及び分母（数２４）はそれぞれ、数１５の分子及び分母に等しいが、これらは解析結果集合にわたる加算ではなく、時刻インデックスにわたる加算によって計算したものである。再推定式（数１５）の分子では、「ｓ_ｉ_ｘ」と「ｏ_ｎ_ｅ」の２個のシーケンスが連続して発生する毎に、各可能な解析結果の尤度が加算される。一方、フォワード・バックワードアルゴリズムを用いた再推定式（数２２ー数２４）においては、「ｓ_ｉ_ｘ」と「ｏ_ｎ_ｅ」の２個のシーケンスが連続して発生し、また、シーケンス「ｏ_ｎ_ｅ」が時刻インデックス（ｔ＋１）で開始するようなすべての解析結果の尤度値をまずグループ化して、加算する。時刻インデックスｔまで加算した時点で加算計算は完了する。

【００６１】上記の例では、「ｓ_ｉ_ｘ」と「ｏ_ｎ_ｅ」の２個のシーケンスが連続して発生し、しかもシーケンス「ｏ_ｎ_ｅ」が時刻インデックス（７）でのみ開始している。ここで、「ｓ_ｉ_ｘ」と「ｏ_ｎ_ｅ」の２個のシーケンスが連続して発生し、また、時刻インデックス（７）でシーケンス「ｏ_ｎ_ｅ」が開始するようなすべての解析結果の尤度値の和は、シーケンス「ｏｎｅｓ_ｉ_ｘｏ_ｎ_ｅｅｉｇｈｔｓｉｘｔｈｒｅｅｔｗｏ」の尤度であり、これは、次式に等しい。

【数２９】

$$\alpha(6,3)\times p(o\_n\_e \mid s\_i\_x)^{(k)}\times\beta(9,3)$$

【００６２】ここで、第２項のｐ（ｏ_ｎ_ｅ｜ｓ_ｉ_ｘ）は、反復回数パラメータ（ｋ）における頻度確率である。また、前方向の変数αの定義により、変数α（６，３）はシーケンス「ｏｎｅｓ_ｉ_ｘ」の尤度であり、さらに、後ろ方向の変数βの定義により、変数β（９，３）は、シーケンス「ｏ_ｎ_ｅ」が得られたときの、シーケンス「ｅｉｇｈｔｓｉｘｔｈｒｅｅｔｗｏ」の尤度である。

【００６３】数１５の分母では、可能な各解析結果の尤度を、シーケンス「ｓ_ｉ_ｘ」がこの解析において発生するのと同じ回数で加算する。等価である、フォワード・バックワードアルゴリズムを用いた前後方向の定式化では、シーケンス「ｓ_ｉ_ｘ」が発生し、時刻インデックス（ｔ）で終了するすべての全解析結果の尤度値をまずグループ化した後に加算し、時刻インデックスｔを越えた時点で加算を終了する。

【００６４】上述の例では、シーケンス「ｓ_ｉ_ｘ」は、時刻インデックス（６）と時刻インデックス（１７）で終了するように発生している。シーケンス「ｓ_ｉ_ｘ」が時刻インデックス（６）で終了するように発生するすべての解析結果の尤度値の加算は、シーケンス「ｏｎｅｓ_ｉ_ｘｏ_ｎ_ｅｅｉｇｈｔｓｉｘｔｈｒｅｅｔｗｏ」の尤度であり、これは次式に等しい。

【００６５】

【数３０】

19

$$\alpha^{(k)}(6,3)\times p(o\_n\_e\mid s\_i\_x)\times\beta(9,3)$$

【００６６】ここで、前方向の変数αの定義により、変数α（６，３）はシーケンス「o ne s_i_x」の尤度であり、後ろ方向の変数βの定義により、変数β（９，３）は、シーケンス「o_n_e」が与えられたときの、シーケンス「ｅｉｇｈｔｓｉｘｔｈｒｅｅｔｗｏ」の尤度であ＊る。

20

【００６７】次いで、時刻インデックス（１７）においてシーケンス「s_i_x」が終了するすべての解析結果の尤度値の加算は、シーケンス「ｏｎｅｓｉｘｏｎｅｅｉｇｈｔ s_i_x t_h_r_e_e ｔｗｏ」の尤度であり、これは次式に等しい。

【００６８】

【数３１】

$$\alpha^{(k)}(17,3)\times p(t\_h\_r\_e\_e\mid s\_i\_x)\times\beta(22,5)$$

【００６９】ここで、前方向の変数αの定義により、変数α（１７，３）はシーケンス「ｏｎｅｓｉｘｏｎｅｅｉｇｈｔ s_i_x」の尤度であり、後ろ方向の変数βの定義により、変数β（２２，５）は、シーケンス「t_h_r_e_e」が与えられたときの、シーケンス「ｔｗｏ」の尤度である。

【００７０】従って、「ｏｎｅｓｉｘｏｎｅｅｅｉｇｈｔｓｉｘｔｈｒｅｅｔｗｏ」なる学習データにおける、反復回数パラメータ（ｋ＋１）における頻度確率p(o_n_e｜s_i_x)に対する、フォワード・バックワードアルゴリズムを用いた再推定式は次式のようになる。

※【００７１】

【数３２】

$$p^{(k+1)}(o\_n\_e\mid s\_i\_x)=p_e/p_t$$

ここで、

【数３３】

$$p_e=\alpha^{(k)}(6,3)\times p(o\_n\_e\mid s\_i\_x)\times\beta(9,3)$$

※【数３４】

$$p_t=\alpha^{(k)}(6,3)\times p(o\_n\_e\mid s\_i\_x)\times\beta(9,3)+\alpha^{(k)}(17,3)\times p(t\_h\_r\_e\_e\mid s\_i\_x)\times\beta(22,5)$$

【００７２】以上説明したように、本発明の実施形態における特徴は、フォワード・バックワードアルゴリズムを用いて、数２３及び数２４を含む数２２を定式化したことにあるが、当該特徴とする数式は、以下の意味を有する。当該式は、入力データにおいて、当該単位列である第２の単位列が第１の単位列に続くときの単位列のシーケンス間のバイグラムの頻度確率を、上記入力データにおける処理対象の各単位列に対して計算するための式であり、上記シーケンス間のバイグラムの頻度確率は、第１と第２の単位列を含むすべてのセグメント化での尤度の和を、第１の単位列を含むすべてのセグメント化での尤度の和で除算することによって得られる。また、上記式は、上記入力データにおいて各単位列が発生する平均回数を示す分母と、上記入力データにおいて第２の単位列が第１の単位列に続くときの各単位列に対する平均回数を示す分子とを有し、上記分子は、処理対象の各単位列に対する、上記前方尤度と、当該単位列の直前の単位列を条件としたときの当該単位列の頻度確率と、上記後方尤度の積の和であり、上記分母は、処理対象の各単位列に対する、上記前方尤度と、当該単位列の直前の単位列を条件としたときのすべての単位列の頻度確率と、上記後方尤度の積の和である。

【００７３】<<Ａ１．３>>前方向の変数αと後ろ方向の変数βの計算例

例として、データ「ｏｎｅｓｉｘｏｎｅｅｉｇｈｔｓｉｘｔｈｒｅｅｔｗｏ」について変数α（９，３）と変数β（９，３）を以下に計算する。ここで、変数α（９，３）は、シーケンス「ｏｎｅｓｉｘｏ_n_e」の尤度であり、このシーケンスは、時刻インデックス９までのシーケンスであって、最後尾において長さ３のシーケンスを有する。また、変数β（９，３）は、シーケンス「o_n_e」が与えられたときの、シーケンス「ｅｉｇｈｔｓｉｘｔｈｒｅｅｔｗｏ」の条件つき尤度であり、このシーケンスは、時刻インデックス９以降のシーケンスであって、先行するシーケンス”o_n_e”は予め知られている。

【００７４】シーケンス”o_n_e”までの尤度（前方の変数）α（９，３）は、次式で計算される。なお、シーケンス（系列）の長さの最大値を”５”に指定した場合について考える。

【数３５】α（９，３）＝下記の加算値

（ａ）n_e_s_i_xについて：α（６，５）×ｐ（o_n_e｜n_e_s_i_x)

（ｂ）e_s_i_xについて：α（６，４）×ｐ（o_n_e｜e_

s_i_x)
（ｃ）s_i_xについて：α（６，３）×ｐ（o_n_e｜s_i_x)
（ｄ）i_xについて：α（６，２）×ｐ（o_n_e｜i_x)
（ｅ）ｘについて：α（６，１）×ｐ（o_n_e｜x)
【００７５】シーケンス”o_n_e”の条件のもとでのその後方の尤度（後方の変数）β（９，３）は、次式で計算される。
【数３６】β（９，３）＝下記の加算値
（ａ）e_i_g_h_tについて：ｐ（e_i_g_h_t｜o_n_e）×β（９＋５，５）（ｂ）e_i_g_hについて：ｐ（e_i_g_h｜o_n_e）×β（９＋４，４）（ｃ）e_i_gについて：ｐ（e_i_g｜o_n_e）×β（９＋３，３）（ｄ）e_iについて：ｐ（e_i｜o_n_e）×β（９＋２，２）
（ｅ）eについて：ｐ（e｜o_n_e）×β（９＋１，１）
【００７６】＜＜Ａ２＞＞クラスの事例
シーケンスがクラスに属するケースでは、上述の例のバイグラムの確率部分を、以下のように置き換えることによって変数α，βが計算される。
（ａ）ｐ（o_n_e｜n_e_s_i_x）は、ｐ（class of o_n_e｜class of n_e_s_i_x）×ｐ（o_n_e｜class of o_n_e）と取って換えられる。
（ｂ）ｐ（o_n_e｜e_s_i_x）は、ｐ（class of o_n_e｜class of e_s_i_x）×ｐ（o_n_e｜class of o_n_e）と取って換えられる。（ｃ）ｐ（o_n_e｜s_i_x）は、ｐ（class of o_n_e｜class of s_i_x）×ｐ（o_n_e｜class of o_n_e）と取って換えられる。
（ｄ）ｐ（o_n_e｜i_x）は、ｐ（class of o_n_e｜class of i_x）×ｐ（o_n_e｜class of o_n_e）と取って換えられる。
（ｅ）ｐ（o_n_e｜x）は、ｐ（class of o_n_e｜class of x）×ｐ（o_n_e｜class of o_n_e）と取って換えられる。
（ｆ）ｐ（e_i_g_h_t｜o_n_e）は、ｐ（class of e_i_g_h_t｜class of o_n_e）×ｐ（e_i_g_h_t｜class of e_i_g_h_t）と取って換えられる。
（ｇ）ｐ（e_i_g_h｜o_n_e）は、ｐ（class of e_i_g_h｜class of o_n_e）×ｐ（e_i_g_h｜class of e_i_g_h）と取って換えられる。
（ｈ）ｐ（e_i_g｜o_n_e）は、ｐ（class of e_i_g｜class of o_n_e）×ｐ（e_i_g｜class of e_i_g）と取って換えられる。
（ｉ）ｐ（e_i｜o_n_e）は、ｐ（class of e_i｜class of o_n_e）×ｐ（e_i｜class of e_i）と取って換えられる。
（ｊ）ｐ（e｜o_n_e）は、ｐ（class of e｜class of o_n_e）×ｐ（e｜class of e）と取って換えられる。
【００７７】＜統計的言語モデル生成処理＞図３は、図１の統計的言語モデル生成部２０によって実行される統計的言語モデル生成処理を示すフローチャートである。ここで、統計的言語モデル生成部２０は、図１に示すように、次のメモリ３１乃至３６に区分されたワーキングＲＡＭ３０を備える。
（ａ）パラメータメモリ３１：当該生成処理で用いる種々の設定パラメータを記憶するメモリである。
（ｂ）シーケンス頻度確率メモリ３２：計算された各シーケンスの頻度確率を記憶するメモリである。
（ｃ）クラス定義メモリ３３：推定された各クラスに属する文字列を記憶するメモリである。
（ｄ）クラス条件付き頻度確率メモリ３４：推定された各クラスに属する各文字列に対する頻度確率、すなわち、クラスの条件付きのクラス間の文字列の頻度確率を記憶するメモリである。
（ｅ）クラスバイグラム頻度確率メモリ３５：クラスのバイグラムの頻度確率を記憶するメモリである。
（ｆ）セグメント化されたシーケンスメモリ３６：再推定処理後のセグメント化されたシーケンス（文字列）を記憶するメモリである。
【００７８】図３において、まず、ステップＳ１では、学習用テキストデータメモリ２１からテキストデータを読み込む。ここで、入力される学習用テキストデータは、離散的な単位のシーケンスであり、ここで、単位とは例えば、文字であり、シーケンスは単語又は文となり得る文字列である。また、予め下記の入力パラメータが設定されてパラメータメモリ３１に記憶されている。
（ａ）シーケンスの最大長（単位の数で表す。）、
（ｂ）再推定処理後のクラス数、（ｃ）廃棄するシーケンス数のしきい値（すなわち、廃棄するシーケンスの発生数の最小値）、及び（ｄ）終了条件。ここで、終了条件は、例えば、反復回数ｋのしきい値である。
【００７９】次いで、ステップＳ２で、初期化処理が実行される。入力された学習用テキストデータにおいて、複数の単位からなるシーケンスの相対的な頻度を計数して、それに基づいて各シーケンスの頻度確率を初期設定する。また、上記設定された廃棄するシーケンス数のしきい値以下のシーケンスについては廃棄する。そして、反復回数パラメータｋを０にリセットする。
【００８０】次いで、ステップＳ３では、ブラウンアルゴリズムを用いた分類処理を実行する。この分類処理では、反復回数パラメータｋのときの各シーケンスの頻度確率に基づいて、クラス間の相互情報量の損失が最小となるように、反復回数パラメータｋのときの、クラス定義、クラス条件付きクラス間のシーケンスの頻度確率、及びクラスバイグラムの頻度確率を計算してそれぞれメモリ３２乃至３５に出力して記憶する。この処理における分類基準は、隣接するシーケンス間の相互情報量であり、上述のアルゴリズムを用いる。これらの相互情報量とアルゴリズムは、隣接する単語の場合に対して、ブラウンによって提案されており、本実施形態では、ブラウンアルゴリズムを用いる。しかしながら、本発明はこれ

に限らず、単位の頻度確率を基礎とする他の分類アルゴリズムを使用することができる。

【００８１】次いで、ステップＳ４において、フォワード・バックワードアルゴリズムを参照して得られた数２２－数２４を用いて、バイーマルチグラムを用いた再推定処理を実行する。この処理では、直前のステップＳ３で計算された、反復回数パラメータｋのときの、クラス定義、クラス条件付きクラス間のシーケンスの頻度確率、及びクラスバイグラムの頻度確率に基づいて、次の反復パラメータのときのシーケンス間のバイグラムの頻度確率の最尤推定値を得るように、反復回数パラメータ（ｋ＋１）のときの、各シーケンスの頻度確率を再推定して計算して、メモリ３２に出力して記憶する。この処理における処理基準は、上記数２２－数２４を用いて、すなわち、複数のシーケンスのクラスとバイグラムの依存性を仮定して計算された解析結果の尤度の中の最大値である最尤推定値を基準値として用いることであり、再推定のためのアルゴリズムとしてＥＭアルゴリズムを用いる。

【００８２】次いで、ステップＳ５で、所定の終了条件を満足するか否かが判断され、ＮＯのときは、ステップＳ６で反復回数パラメータｋを１だけインクリメントしてステップＳ３及びＳ４の処理を繰り返す。一方、ステップＳ５でＹＥＳであれば、生成された統計的言語モデルのデータを統計的言語モデルメモリ２２に出力して記憶する。ここで、生成された統計的言語モデルのデータとは、各シーケンスの頻度確率に関するデータであり、具体的には、下記のデータである。

（ａ）入力されたデータを複数のシーケンスにセグメント化したときの最尤推定値を有する各シーケンスのデータ；

（ｂ）クラス定義、すなわち、各クラスにおけるシーケンス；及び

（ｃ）クラスの頻度確率、すなわち、各クラスのバイグラム確率、各シーケンスのクラス条件付き確率。

【００８３】図４は、図３のサブルーチンであるブラウンアルゴリズムを用いた分類処理を示すフローチャートである。単語の自動分類のために、ブラウン他によってシーケンスの自動分類に使用するためのアルゴリズム（例えば、従来技術文献５参照。）が提案されており、本実施形態では、これを使用する。ブラウンらは、文章の尤度を最大化するクラスへの分割又はセグメント化

が、隣接する単語間の相互情報量を最大化する分割又はセグメント化でもあることを示している。彼らは単語のバイグラム分布を入力とし、単語クラスへの分割及びクラス分布を出力する貪欲なアルゴリズム(greedy algorithm)を提案している。一方、本発明者は、入力としてバイーマルチグラムの頻度確率の分布（すなわち、シーケンスのバイグラムの頻度確率の分布）を採用することにより、このアルゴリズムを適用している。出力は、シーケンスのクラスへのセグメント化及びその各シーケンスの頻度確率の分布である。

【００８４】この分類処理で用いる相互情報量を用いた単語のクラスタリングについて詳細説明する（例えば、従来技術文献６「北研二ほか著，”音声言語処理”，森北出版，ｐｐ．１１０－１１３，１９９６年１１月１５日発行」参照。）。ここでは、隣接する単語に基づく単語の分類法として、クラス間の相互情報量を最大にする方法について説明する。相互情報量に基づくクラスタリングは、バイグラムのクラスモデルにおいて単語をクラスへ分割する最尤な方法は、隣接するクラスの平均相互情報量を最大にするようなクラス割り当てであることを、理論的な根拠としている。Ｎ－ｇｒａｍのクラスモデルとは、次式のように、単語のクラスのＮ－ｇｒａｍとクラス別の単語の出現分布の組み合わせで、単語のＮ－ｇｒａｍを近似する言語モデルのことである(この式は、単語クラスを品詞に置き換えれば、形態素解析におけるHMMの式と同じになる。従って、この単語分類法は、最適な品詞体系を自動的に求める方法とも考えられる。

【数３７】$P(w_i \mid w_1^{i-1}) \fallingdotseq P(w_i \mid c_i)\, P(c_i \mid c_{i-n+1}^{i-1})$

【００８５】ここで、単語w_iをクラスc_iに写像する関数πを用いて、Ｖ個の単語をＣ個のクラスに分割すると仮定する。学習テキストt_1^Tが与えられたとき、$P(t_2^T \mid t_1) = P(T_2 \mid T_1)\, P(t_3 \mid t_2) \cdots P(t_T \mid t_{T-1})$を最大にするように関数$\pi$を決めればよい。詳細は省略するが、単語あたりの対数尤度Ｌ（π）、単語のエントロピーＨ（ｗ）、隣接するクラスの平均相互情報量Ｉ（c_1；c_2）の間には、近似的に次式の関係が成り立つ。

【００８６】

【数３８】

25 26

$$
\begin{aligned}
&L(\pi)\\
&=(T-1)^{-1}\log P(t_2^T \mid t_1)\\
&=\sum_{w_1,w_2}\{C(w_1w_2)\}/(T-1)\times\log P(c_2 \mid c_1)\,P(w_2 \mid c_2)\\
&\fallingdotseq\sum_{c_1,c_2}P(c_1c_2)\log\{P(c_2 \mid c_1)/P(c_2)\}\\
&\quad+\sum_{w}P(w)\log P(w)\\
&=I(c_1;c_2)-H(w)
\end{aligned}
$$

【００８７】ここで、Ｈ（ｗ）は分割πに依存しないから、Ｌ（π）を最大化するためには、Ｉ（c_1；c_2）を最大化すればよい。いまのところ、平均相互情報量を最大化するような分割を求めるアルゴリズムは知られていない。しかしながら、本実施形態で用いる次のような貪欲なアルゴリズム（greedy algorithm）でも、かなり興味深いクラスタを得ることができる。このように包含関係を持つクラスタを生成する方法は、階層的クラスタリングと呼ばれる。これに対して、ｋ平均アルゴリズムのように、重なりを持たないクラスタを生成する方法は非階層的クラスタリングと呼ばれる。

【００８８】次の併合をＶ－１回繰り返すと、すべての単語が一つのクラスになる。すなわち、クラスが併合される順序から、単語を葉とする二分木ができる。

１．すべての単語に対して、一つのクラスを割り当てる。

２．可能な二つのクラスの組み合わせの中で、平均相互情報量の損失を最小にする組み合わせを選択し、これらを一つのクラスに併合する。

３．ステップ２をＶ－Ｃ回繰り返すとＣ個のクラスが得られる。

【００８９】一般に、クラスタが形成される過程を表す階層構造は樹形図（dendrogram）と呼ばれるが、自然言語処理ではこれをシソーラスの代わりに使うことができる。単純に考えると、この準最適なアルゴリズムは、語彙数Ｖに対してV^5の計算量を必要とする。しかし、（１）二つのクラスタを併合したときの情報量の変化だけを求めればよいことや、（２）二つのクラスタの併合により相互情報量が変化するのは全体の一部に過ぎないことを利用すれば、Ｏ（V^3）の計算、すなわち、繰り返し回数Ｖの三乗に比例するオーダーの計算コストで済む。

【００９０】分類処理（又はクラスタリング処理）を示す図４において、まず、ステップＳ１１では、初期設定処理が実行され、各シーケンスをその自らのクラスに割り当てる。すなわち、各シーケンスs_iそれぞれ各クラスC_iに割り当てる。従って、クラスの初期バイグラムの頻度確率の分布はシーケンスのバイグラムの頻度確率の分布に等しく、また、

【数３９】$p(s_i \mid C_i)=1$

である。

【００９１】次いで、ステップＳ１２で、各クラスの対（C_k，C_l）について、クラスC_kとクラスC_lとをマージしたときの相互情報量の損失を計算した後、ステップＳ１３で、相互情報量の損失が最小であるクラスの対をマージする。そして、ステップＳ１４で、上記マージに従って、メモリ３４及び３５に記憶されたクラスの頻度確率の分布を更新する。次いで、ステップＳ１５で、ステップＳ２の初期化処理で設定された必要なクラス数が得られたか否かが判断され、ＮＯであるときは、ステップＳ１２に戻り、上記の処理を繰り返す。一方、ステップＳ１５で、ＹＥＳのときは、元のメインルーチンに戻る。

【００９２】＜音声認識装置＞次いで、図１に示す連続音声認識装置の構成及び動作について説明する。図１において、単語照合部４に接続された音素隠れマルコフモデル（以下、隠れマルコフモデルをＨＭＭという。）メモリ１１内の音素ＨＭＭは、各状態を含んで表され、各状態はそれぞれ以下の情報を有する。

（ａ）状態番号、（ｂ）受理可能なコンテキストクラス、（ｃ）先行状態、及び後続状態のリスト、（ｄ）出力確率密度分布のパラメータ、及び（ｅ）自己遷移確率及び後続状態への遷移確率。なお、本実施形態において用いる音素ＨＭＭは、各分布がどの話者に由来するかを特定する必要があるため、所定の話者混合ＨＭＭを変換して生成する。ここで、出力確率密度関数は３４次元の対角共分散行列をもつ混合ガウス分布である。また、単語照合部４に接続された単語辞書メモリ１２内の単語辞書は、音素ＨＭＭメモリ１１内の音素ＨＭＭの各単語毎にシンボルで表した読みを示すシンボル列を格納する。

【００９３】図１において、話者の発声音声はマイクロホン１に入力されて音声信号に変換された後、特徴抽出部２に入力される。特徴抽出部２は、入力された音声信号をＡ／Ｄ変換した後、例えばＬＰＣ分析を実行し、対数パワー、１６次ケプストラム係数、Δ対数パワー及び１６次Δケプストラム係数を含む３４次元の特徴パラメータを抽出する。抽出された特徴パラメータの時系列はバッファメモリ３を介して単語照合部４に入力される。

【００９４】単語照合部４は、ワンーパス・ビタビ復号化法を用いて、バッファメモリ３を介して入力される特

徴パラメータのデータに基づいて、音素HMM１１と単語辞書１２とを用いて単語仮説を検出し尤度を計算して出力する。ここで、単語照合部４は、各時刻の各HMMの状態毎に、単語内の尤度と発声開始からの尤度を計算する。尤度は、単語の識別番号、単語の開始時刻、先行単語の違い毎に個別にもつ。また、計算処理量の削減のために、音素HMM１１及び単語辞書１２とに基づいて計算される総尤度のうちの低い尤度のグリッド仮説を削減する。単語照合部４は、その結果の単語仮説と尤度の情報を発声開始時刻からの時間情報（具体的には、例えばフレーム番号）とともにバッファメモリ５を介して単語仮説絞込部６に出力する。

【００９５】単語仮説絞込部６は、単語照合部４からバッファメモリ５を介して出力される単語仮説に基づいて、統計的言語モデルメモリ２２内の統計的言語モデルを参照して、終了時刻が等しく開始時刻が異なる同一の単語の単語仮説に対して、当該単語の先頭音素環境毎に、発声開始時刻から当該単語の終了時刻に至る計算された総尤度のうちの最も高い尤度を有する１つの単語仮説で代表させるように単語仮説の絞り込みを行った後、絞り込み後のすべての単語仮説の単語列のうち、最大の総尤度を有する仮説の単語列を認識結果として出力する。本実施形態においては、好ましくは、処理すべき当該単語の先頭音素環境とは、当該単語より先行する単語仮説の最終音素と、当該単語の単語仮説の最初の２つの音素とを含む３つの音素並びをいう。

【００９６】例えば、図２に示すように、（ｉ－１）番目の単語Ｗｉ－１の次に、音素列ａ１，ａ２，…，ａｎからなるｉ番目の単語Ｗｉがくるときに、単語Ｗｉ－１の単語仮説として６つの仮説Ｗａ，Ｗｂ，Ｗｃ，Ｗｄ，Ｗｅ，Ｗｆが存在している。ここで、前者３つの単語仮説Ｗａ，Ｗｂ，Ｗｃの最終音素は／ｘ／であるとし、後者３つの単語仮説Ｗｄ，Ｗｅ，Ｗｆの最終音素は／ｙ／であるとする。終了時刻ｔｅと先頭音素環境が等しい仮説（図２では先頭音素環境が“ｘ／ａ１／ａ２”である上から３つの単語仮説）のうち総尤度が最も高い仮説（例えば、図２において１番上の仮説）以外を削除する。なお、上から４番めの仮説は先頭音素環境が違うため、すなわち、先行する単語仮説の最終音素がｘではなくｙであるので、上から４番めの仮説を削除しない。すなわち、先行する単語仮説の最終音素毎に１つのみ仮説を残す。図２の例では、最終音素／ｘ／に対して１つの仮説を残し、最終音素／ｙ／に対して１つの仮説を残す。

【００９７】以上の実施形態においては、当該単語の先＊頭音素環境とは、当該単語より先行する単語仮説の最終音素と、当該単語の単語仮説の最初の２つの音素とを含む３つの音素並びとして定義されているが、本発明はこれに限らず、先行する単語仮説の最終音素と、最終音素と連続する先行する単語仮説の少なくとも１つの音素とを含む先行単語仮説の音素列と、当該単語の単語仮説の最初の音素を含む音素列とを含む音素並びとしてもよい。

【００９８】以上の実施形態において、特徴抽出部２と、単語照合部４と、単語仮説絞込部６と、統計的言語モデル生成部２０とは、例えば、デジタル電子計算機などのコンピュータで構成され、バッファメモリ３，５と、音素HMMメモリ１１と、単語辞書メモリ１２と、学習用テキストデータメモリ２１と、統計的言語モデルメモリ２２とは、例えばハードデイスクメモリなどの記憶装置で構成される。

【００９９】以上実施形態においては、単語照合部４と単語仮説絞込部６とを用いて音声認識を行っているが、本発明はこれに限らず、例えば、音素HMM１１を参照する音素照合部と、例えばＯｎｅ　Ｐａｓｓ　ＤＰアルゴリズムを用いて統計的言語モデルを参照して単語の音声認識を行う音声認識部とで構成してもよい。

【０１００】

【実施例】＜統計的言語モデル生成処理の第１の実施例＞入力される学習データが、以下のような１０００文字列の場合であり、単位である文字から単語にセグメント化するための例である。

「ｏｎｅｓｉｘｏｎｅｅｉｇｈｔｆｉｖｅｚｅｒｏ...」

但し、奇数の単語の後には必ず偶数の単語が後続し、偶数の単語の後には必ず奇数の単語が後続する場合である。当該実施例における入力パラメータは以下の通りである。

（ａ）１個のシーケンスの最大長＝５、（ｂ）クラス数＝２、及び（ｃ）廃棄するシーケンスのしきい値＝１００。

【０１０１】初期化処理（ｋ＝０）では、学習データにおいて、１００回を越えて観測した文字のすべての組合せの相対的な計数値を初期値とする。従って、反復パラメータｋ＝０におけるシーケンスの頻度確率の分布の計数結果は次の表のようになる。なお、各シーケンスのｎｂ（・）は計数値を表す。

【０１０２】

【表２】

Ｐ（ｎ｜ｏ）＝ｎｂ（ｏｎ）／ｎｂ（ｏ）＝０．０８	
ｐ（ｎ_ｅ｜ｏ）＝ｎｂ（ｏｎｅ）／ｎｂ（ｏ）＝０．０６	
・・・	
ｐ（ｎ_ｅ_ｓ_ｉ_ｘ｜ｏ）＝ｎｂ（ｏｎｅｓｉｘ）／ｎｂ（ｏ）＝０．００５	

29　　30

p (e | o_n) ＝ｎｂ (one) ／ｎｂ (on) ＝０．９
p (e_s | o_n) ＝ｎｂ (ones) ／ｎｂ (on) ＝０．００５
．．．
p (e_s_i_x_o | o_n) ＝ｎｂ (onesixo) ／ｎｂ (on) ＝０．００１
．．．
p (s_i_x | o_n_e) ＝ｎｂ (onesix) ／ｎｂ (one) ＝０．０５
．．．

【０１０３】ステップＳ３の分類処理では、入力データは、反復パラメータｋ＝０のときのシーケンスの頻度確率の分布であり、当該分類処理における出力データは、以下のようになる。
（ａ）反復パラメータｋ＝１のときのクラス定義
【数４０】ｃｌａｓｓ１＝｛e_s_i_x_o;e;e_t_w_o;n_e_s_i_x;.....;f_o_u_r;f_o_u_r_f;...;g_h_t_s;g_h_t_o_n_e;e_i_g_h_t｝
【数４１】ｃｌａｓｓ２＝｛o_n_e;e_s_i_x_o;x;f_i_v;f_i_v_e;t_s_e_v;s_e_v_e_n;.....;x_n_i;x_n_i_n_e;n_i_n_e;...｝
ｃｌａｓｓ３＝……
（ｂ）反復パラメータｋ＝１のときのクラス条件付き頻度確率の分布
【数４２】
p (e_s_i_x_o | class 1) , p (e | class 1) , . . .
p (o_n_e | class 2) , p (e_s_i_x_o | class 2) , . . .
（ｃ）反復パラメータｋ＝１のときのクラスバイグラムの頻度確率の分布
【数４３】p (class 1 | class 2) ＝０．３
p (class 2 | class 1) ＝０．１
p (class 3 | class 1) ＝０．４
．．．
【０１０４】ステップＳ４の再推定処理では、反復パラメータｋ＝１のときのクラス定義及びクラスの頻度確率の分布を入力データとし、次に示す反復パラメータｋ＝１のときのシーケンスの頻度確率の分布を出力する。
【数４４】p (n | o) ＝０．９
p (n_e | o) ＝０．８
p (n_e_s | o) ＝０．０５
．．．
p (n_e_s_i_x | o) ＝０
【数４５】p (e | o_n) ＝０．０２
p (e_s | o_n) ＝０．００１
．．．
p (e_s_i_x_o | o_n) ＝０
．．．
p (s_i_x | o_n_e) ＝０．５
．．．
【０１０５】以下同様に処理が実行され、第１の実施例における出力結果は以下のようになる。
（ａ）セグメント化された入力文字列（ＭＬセグメント化）
”o_n_e s_i_x o_n_e e_i_g_h_t f_i_v_e z_e_r_o ...”
（ｂ）クラス定義
【数４６】ｃｌａｓｓ１＝｛o_n_e;t_h_r_e_e;f_i_v_e;s_e_v_e_n;n_i_n_e｝
ｃｌａｓｓ２＝｛z_e_r_o;t_w_o;f_o_u_r;s_i_x;e_i_g_h_t｝
（ｃ）クラス条件付きの頻度確率の分布
【数４７】p (o_n_e | class 1) ＝０．２
p (t_h_r_e_e | class 1) ＝０．２
p (f_i_v_e | class 1) ＝０．２
．．．
p (z_e_r_o | class 2) ＝０．２
p (t_w_o | class 2) ＝０．２
（ｄ）クラスバイグラムの頻度確率の分布
【数４８】p (class 1 | class 2) ＝１
p (class 2 | class 1) ＝１
【０１０６】＜統計的言語モデル生成処理の第２の実施例＞入力される学習データが、自然言語のテキストデータによる以下の文、すなわち単語列である場合であって、単位である単語をフレーズにセグメント化する場合を説明するための実施例である。ここで、＜ｓ＞は開始を示す記号であり、＜／ｓ＞は終了を示す記号である。
「<s> good afternoon new washington hotel may i help you ...</s>」
ここで、入力パラメータは、以下の通りである。
（ａ）シーケンスの最大長＝数個の単語（例えば、１乃至５個の単語、以下の実施例では、４）、（ｂ）クラス数＝１０００、及び（ｃ）初期化処理のしきい値＝３０。
【０１０７】初期化処理（ｋ＝０）では、学習データにおいて、３０回を越えて観測した単語のすべての組合せの相対的な計数値を初期値とする。従って、反復パラメータｋ＝０におけるシーケンスの頻度確率の分布の計数結果は次の表のようになる。
【０１０８】
【表３】

31　　32

```
  p (afternoon | good)
=nb (good afternoon) ／nb (good) =0．08
  p (afternoon_new | good)
=nb (good afternoon new) ／nb (good) =0．06
  p (good_afternoon | ⟨s⟩)
=nb (⟨s⟩good afternoon) ／nb (⟨s⟩) =0．06
...
  p (⟨/s⟩ | may_i_help_you)
=nb (may i help you ⟨/s⟩) ／nb (may i help you)
=0．005
```

【０１０９】そして、第２の実施例における出力結果は以下のようになる。

（ａ）セグメント化された入力文字列（ＭＬセグメント化）

「good_afternoon new_washington_hotel may_i_help_you」

（ｂ）クラス定義

【数４９】ｃｌａｓｓ１＝｛good_afternoon ; good_morning;hello ; may_i_help_you...｝

...

ｃｌａｓｓ２＝｛new_washington_hotel ; sheraton_hotel ; plaza;...｝

...

ｃｌａｓｓ１０００＝｛give_me_some ; tell_me｝

（ｃ）クラス条件付き頻度確率の分布

【数５０】

p（good_afternoon｜class 1）＝０．００３

p（good_morning｜class 1）＝０．００２

p（hello｜class 1）＝０．００２

...

（ｄ）クラスバイグラムの頻度確率の分布

【数５１】p（class 2｜class 1）＝０．０４

p（class 3｜class 1）＝０．００５

...

【０１１０】＜実験及び実験結果＞本発明者は、実施形態の装置の性能を実験するために、下記の実験を行った。まず、プロトコル及びデータベースの実験及び実験結果について述べる。可変長フレーズ間のバイグラム依存を学習する目的は、従来のワードバイグラムモデルの限界を改善する一方で、モデル内のパラメータ数を単語のトライグラムの場合よりも少なくすることにある。従って、バイーマルチグラムモデルの評価を行うために適する基準は、その予測能力、パラメータ数を測定し、従来のバイグラム、トライグラムモデルのそれらと比較することである。予測能力は通常、次式のパープレキシティの測定によって評価される。

【０１１１】

【数５２】

ＰＰ＝ｅｘｐ｛－（１／Ｔ）ｌｏｇ（Ｌ（Ｗ））｝

【０１１２】ここで、Ｔを文Ｗにおける単語の数である。パープレキシティＰＰが低いほど、モデルの予測がより高精度であることを示す。統計的モデルでは、実際には２つのパープレキシティ値ＰＰ及びPP^*が存在し、数５２の中のＬ（Ｗ）をそれぞれ次式として計算される。

【０１１３】

【数５３】

$$L(W) = \sum_{S} L(W, S)$$

及び

【数５４】$L(W) = L(W, S^*)$

【０１１４】２つのパープレキシティPP^*－ＰＰの差は、常に正の数又は零であり、文Ｗの解析結果Ｓの曖昧さの度合い、あるいは発話認識機のように最良の解析結果の尤度を用いて文の尤度に到達する場合は、予測の正確さにおける損失を測定する。

【０１１５】以下では、先ず、ある推定手順における損失（PP^*－ＰＰ）を評価し、この推定手順自体の影響力についてフォワード・バックワードアルゴリズム（数１５）又は決定論的方法（数１６）を用いて考察する。最後に、これら結果を従来のｎ－ｇｒａｍモデルを用いて得られた結果と比較する。本目的の達成のため、クラークソン（（Clarkson）ほか1997年）による公知のＣＭＵツールキットを用いる。実験対象として、次の表の本特許出願人が所有する「旅行の手配」に関するデータを使用する。

【０１１６】

【表４】

本特許出願人が所有する「旅行の手配」に関するデータ

	学習	テスト
文の数	１３６５０	２４３０
トークンの数	１６７０００	２９０００（1%OOV）
語彙数	３５２５	＋２８０OOV

（注）OOVは、Ｏｕｔ　Ｏｆ　Ｖｏｃａｂｕｌａｒｙの略であり、語彙にない単語をいう。

【０１１７】本データベースは、ホテルのクラークと顧客の間で自発的に行われた旅行／宿泊施設情報についての対話である。言いよどみの単語、及び間違った開始は、単一のマーカー“*ｕｈ*”にマッピングされる。本実験において、フレーズの最大長はｎ＝１語から４語まで変化させた（ｎ＝１ではバイーマルチグラムは従来のバイグラムに相当する）。すべてのバイーマルチグラムの頻度確率は、６回のトレーニング反復で推定され、初期化において２０回以下、各反復において１０回以下の頻度でしか現れないすべての文を放棄し、フレーズ辞書の枝刈りを行った。ここで、初期化におけるしきい値が１０－３０の範囲にあるとき、本データにおいて、異なる枝刈り限界値を用いても結果に重大な影響が及ぶことはない。反復の場合のしきい値はその約半分である。

【０１１８】しかしながら、すべての1単語フレーズは、その推定出現回数にかかわらず維持されるため（フレーズs_i及びs_jが１単語フレーズであり、組み合わせｃ（s_i，s_j）の再推定値が零であると、組み合わせｃ（s_i，s_j）は1にリセットされる。）、すべてのワードバイグラムが最終辞書に現れることになる。さらに、すべてのｎ－ｇｒａｍ及びフレーズのバイグラム確率は、ウィッテン（Witten）ほか（１９９１年）による公知のWitten-Bellディスカウンティング法を用いて、カッツ（Katz）（１９８７年）による公知のバックオフ・スムージング法で平滑化される。ここで、Witten-Bellディスカウンティング法を選択したのは、本テストデータにおいて従来のｎ－ｇｒａｍを用いた場合、最良のパープレキシティスコアが得られるためである。

【０１１９】次いで、クラスタリングを行わない実験について述べる。まず、非決定性の方式の度合いにおいては、表４の本特許出願人が所有する「旅行の手配」に関するデータに対するテストで、フォワード・バックワードアルゴリズムによる学習の後に得られたパープレキシティ値ＰＰ*及びＰＰを次の表に示す。パープレキシティ値の差（ＰＰ*－ＰＰ）は通常、パープレキシティの約１ポイント以内にとどまる。すなわち、単一の最良フレーズに依存しても、予測の正確さが大幅に損なわれることがあってはならないことを意味している。

【０１２０】

【表５】

非決定性の方式の度合い

n	1	2	3	4
ＰＰ	５６．０	４３．９	４４．２	４５．０
ＰＰ*	５６．０	４５．１	４５．４	４６．３

【０１２１】次いで、再推定手順の影響力では、フォワード・バックワードアルゴリズム又はビタビ推定アルゴリズムのいずれかを用いたパープレキシティ値ＰＰ*及びモデルサイズを次の表に示す。

【０１２２】

【表６】

推定方法の影響：テストパープレキシティ値　ＰＰ*

n	1	2	3	4
ＦＢ法	５６．０	４５．１	４５．４	４６．３
ビタビ法	５６．０	４５．７	４５．９	４６．２

【０１２３】

★　★【表７】

推定方法の影響：モデルのサイズ

n	1	2	3	4

35　　36

ＦＢ法	３２５０５	４４３８２	４３６７２	４３１８６
ビタビ法	３２５０５	６５１４１	６７２５８	６７２９５

【０１２４】表６及び表７から明らかなように、パープレキシティ値に関する限り、推定方法はほとんど影響を及ぼさず、フォワード・バックワードアルゴリズムによる学習を用いる方がわずかながら有利であるように見える。一方、モデルのサイズは、学習終了時に個々のバイーマルチグラム数として測定された場合、フォワード・バックワードアルゴリズムによる学習において約３０％も減少する。すなわち、同じテストパープレキシティ値に対して、おおよそ４０，０００対６０，０００の違いとなる。

【０１２５】バイーマルチグラム結果は、概して、フレーズ放棄を行う枝刈りのための発見的知識では完全に過学習を回避できないことを示唆する。確かに、（おそら＊くは６から８語にまたがる依存性を意味する）ｎ＝３，４のパープレキシティ値は、（依存性が４語に限定される）ｎ＝２のときのそれよりも高くなる。他の方法、おそらくは短いものよりも長いフレーズを不利にするような方法であれば成功ものと考えられる。

【０１２６】さらに、ｎ－ｇｒａｍとの比較においては、フォワード・バックワードアルゴリズムによる学習から得られたパープレキシティ値（ＰＰ）、ｎ－ｇｒａｍに対するモデルサイズ、及びバイーマルチグラムを次の表に示す。

【０１２７】

【表８】

ｎ－ｇｒａｍの比較

テストパープレキシティ値　ＰＰ				
ｎの値	1	2	3	4
ｎ－ｇｒａｍ	３１４．２	５６．０	４０．４	３９．８
バイーマルチグラム	５６．０	４３．９	４４．２	４５．０

【０１２８】　※　※【表９】

ｎ－ｇｒａｍの比較

モデルのサイズ				
ｎ値	1	2	3	4
ｎ－ｇｒａｍ	３５２６	３２５０５	７５５１１	１１２１４８
バイーマルチグラム	３２５０５	４４３８２	４３６７２	４３１８６

【０１２９】表８及び表９から明らかなように、最も低いバイーマルチグラムパープレキシティスコア（４３．９）は、トライグラムの値よりも依然として高いが、バイグラム値（５６．０）よりもトライグラム値（４０．４）により近い値となっている。さらに、トライグラムスコアはディスカウントされた方法に依存する。なお、線形ディスカウンティング法では、本テストにおけるトライグラムのパープレキシティは、４８．１であった。

【０１３０】５－グラムのパープレキシティ値（上記表に示さず）は４０．８であり、４－ｇｒａｍスコアよりもやや高い。これは、バイーマルチグラムパープレキシティがｎ＞２（すなわち、依存性が４語以上にわたる場合）のとき減少しないという事実に一致する。最後に、バイーマルチグラムモデルのエントリ数はトライグラムモデルのエントリ数よりも少なく（４５０００に対して７５０００）、マルチグラムが達成するモデルの正確性とモデルサイズ間のトレードオフが示されている。

【０１３１】さらに、クラスタリングを用いた実験及び実験結果について述べる。本実験では、フレーズのクラスタリングによってパープレキシティスコアは改善されなかった。パープレキシティの増加が非常に少なくなる（１ポイント以下）のは、フレーズのほんの一部（１０～２０％）のみがクラスタとなる時であり、これを越えるとパープレキシティはかなり悪化する。この効果は、クラス推定が単語推定に統合されない時、ｎ－ｇｒａｍの枠組みにおいても度々報告されている。しかしながら、フレーズのクラスタリングによって、自然発話を特徴づける言いよどみの語の挿入等、ことばの非流暢性のいくつかを自然に扱うことができる。この点を説明するために、先ずｎ＝４語までのフレーズを扱うモデルの学

習の間に統合されるフレーズを次の表に列挙する。ここで、言いよどみを示す“*ｕｈ*”を含むフレーズはこの表の上部に示す。主に、話者の言いよどみによるフレー＊ズの違いは、共に統合されることが多い。

【０１３２】

【表１０】

４語シーケンスまでを扱うモデルにおける統合されたフレーズの一例

{yes_that_will;*uh*_that_would}
{yes_that_will_be;*uh*_yes_that's}
{*uh*_by_the;and_by_the}
{yes_*uh*_i;i_see_i}
{okay_i_understand;*uh*_yes_please}
{could_you_recommend;*uh*_is_there}
{*uh*_could_you_tell;and_could_you_tell}
{so_that_will;yes_that_will;yes_that_would;uh*_that_would}
{if_possible_i'd_like;we_would_like;*uh*_i_want}
{that_sounds_good;*uh*_i_understand}
{*uh*_i_really;*uh*_i_don't}
{*uh*_i'm_staying;and_i'm_staying}
{all_right_we;*uh*_yes_i}

{good_morning_this;good_afternoon_this}
{yes_i_do;yes_thank_you}
{we'll_be_looking_forward;we_look_forward}
{dollars_a_night;and_forty_yen}
{for_your_help;for_your_information}
{hold_the_line;want_for_a_moment}
{yes_that_will_be;and_could_you_tell}
{please_go_ahead;you_like_to_know}
{want_time_would_you;and_you_would}
{yes_there_is;but_there_is}
{join_phillips_in_room;ms._suzuki_in}
{name_is_suzuki;name_is_ms._suzuki}
{i'm_calling_from;a;also_i'd_like}
{much_does_it_cost;can_reach_you}
{thousand_yen_room;dollars_per_person}
{yes_i_do;yes_thank_you;i_see_sir}
{you_tell_me_where;you_tell_me_what}
{a_reservation_for_the;the_reservation_for}
{your_name_and_the;you_give_me_the}
{amy_harris_in;is_amy_harris_in}
{name_is_mary_phillips;name_is_kazuo_suzuki}
{hold_on_a_moment;wait_a_moment}
{give_me_some;also_tell_me}

【０１３３】カワハラ（Kawahara）ら（１９９７年）によれば、上記の表はさらに、単語予測とは別に、フレーズ検索及びクラスタリングを行う他の動機づけ、すなわちトピックの識別や対話のモデリング、及び言語理解に関する問題への対応を示している。確かに本実験におけるクラスタとなったフレーズは、完全盲目的、すなわち意味論的／語用論的情報を全くなくして導かれたものであるが、クラス内フレーズには強固な意味論的相関関係が示されている。しかしながら、本手法を音声理解に効率的に使用できるようにするためには、拘束条件は、例えばスピーチアクトタグ（speech act tags）のようないくつかのより高いレベルの情報を用いてフレーズクラスタリング処理に設定する必要がある。

【０１３４】以上説明したように、フレーズ間にｎ－ｇ

ｒａｍ依存を仮定する可変長フレーズを導くアルゴリズムは、言語モデリングのタスクのために提案され、推定されてきた。特定タスクの言語コーパスは、文をフレーズに構成することによりバイグラムパープレキシティ値を大幅に減らし、一方で言語モデルにおけるエントリ数をトライグラムモデルの場合に比べてより低い値に保つことが可能であることを示している。しかしながら、これら結果は、より効率的な枝刈り方法によってさらに改善され、不要な学習を行わずにより長い依存性について学習することが可能となる。さらに、語形変化の態様を簡単に本枠組み内に統合することができるため、異なる長さを有するフレーズに共通のラベルを割り当てることが可能である。フレーズの意味論的関係が統合されるので、本手法は対話モデリングや言語理解の分野においても用いられる。その場合、意味論的／語用論的情報を用いれば、フレーズクラスを得るための処理に制限を設けることができる。

【０１３５】＜変形例＞以上の実施形態においては、単位は英語の文字であり、シーケンスは単語であり、上記分類処理は、文字列を複数の単語の列に分類し、上記統計的シーケンスモデルは、統計的言語モデルである。本発明はこれに限らず、単位は、日本語などの他の自然言語の文字であってもよい。また、単位は自然言語の単語であり、シーケンスはフレーズであり、上記分類処理は、単語列を複数のフレーズの列に分類し、上記統計的シーケンスモデルは、統計的言語モデルであってもよい。

【０１３６】＜実施形態の効果＞以上説明したように，本発明に係る実施形態によれば、以下のような特有の効果を有する。

（Ａ）ＥＭアルゴリズムを使用して単語のシーケンスの頻度分布を計算することができ、ＭＬ基準を最適化することができる。すなわち、本実施形態のアルゴリズムを用いられば、必ず、クラスタリングの処理を単調収束させることができて、最適値の解析結果を得ることができる。

（Ｂ）シーケンス分類の解析を自由にすることができる。具体的には、上述のフォワード・バックワードアルゴリズムを用いた非決定性の手法を用いるので、自由度のある解が得られる。なお、当該非決定性の手法を用いることができるのは、変数α，βを決めることができるからである。従って、入力データの尤度を改善することにより、シーケンス［ｂｃｄ］が入力シーケンスにあったときに、［ｂｃ］＋［ｄ］、［ｂ］＋［ｃｄ］、［ｂ］＋［ｃ］＋［ｄ］等の小シーケンスへの分割が可能である。言い換えれば、あるシーケンスが入力シーケンスに与えられていても、解析は事前に決定されず、すべては入力データの尤度に依存する、つまり確定的ではなく、入力データの頻度確率に依存してクラスタリングの処理が行われる。

（Ｃ）可変長のシーケンスの自動的分類を行うことができる。ここで、シーケンスの分類を、単語の分類に依存させない。また、シーケンスの分類を直接的に自動的に行なって、長さの違う共通のクラスシーケンスに高精度で分類できる。

【０１３７】従って、本発明に係る実施形態によれば、従来例に比較して、最適な状態に向かう単調な収束を保証することができ、自由度があり、可変長のシーケンスを同一のクラスで取り扱うことができ、ディジタル計算機を用いて実用的に高速処理することができる統計的シーケンスモデル生成装置、統計的言語モデル生成装置及び音声認識装置を提供することができる。

【０１３８】

【発明の効果】以上詳述したように本発明に係る統計的シーケンスモデル生成装置によれば、１個又は複数の単位からなる単位列であるシーケンスを含む入力データに基づいて、可変長の自然数N_1個の単位列であるマルチグラムと、可変長の自然数N_2個の単位列であるマルチグラムとの間のバイグラムであるバイーマルチグラムの統計的シーケンスモデルを生成する統計的シーケンスモデル生成装置であって、上記入力データに基づいて、予め決められたN_1，N_2の最大値の拘束条件のもとで、すべての単位列の組み合わせの上記バイグラムの頻度確率を計数する初期化手段と、上記初期化手段によって計数された上記バイグラムの頻度確率に基づいて、各クラスの対をマージしたときの相互情報量の損失が最小となるようにマージして各クラスの頻度確率を更新して予め決められた数の複数のクラスに分類することにより、分類されたクラスに含まれる単位列と、分類されたクラスの条件付きの単位列の頻度確率と、分類されたクラス間のバイグラムの頻度確率を計算して出力する分類手段と、上記分類処理手段から出力される分類されたクラスに含まれる単位列と、分類されたクラスの条件付きの単位列の頻度確率と、分類されたクラス間のバイグラムの頻度確率とに基づいて、ＥＭアルゴリズムを用いて、最尤推定値を得るように再推定し、ここで、フォワード・バックワードアルゴリズムを用いて、処理対象の各単位列に対して、時系列的に前方にとり得る処理対象の当該単位列に対する前方尤度と、当該単位列の直前の単位列を条件としたときの当該単位列の頻度確率と、時系列的に後方にとり得る当該単位列に対する後方尤度とに基づいてシーケンス間のバイグラムの頻度確率を示す式を用いて、当該シーケンス間のバイグラムの頻度確率を再推定することにより、再推定結果である上記バイーマルチグラムの統計的シーケンスモデルを生成して出力する再推定手段と、上記分類手段の処理と上記再推定手段の処理を所定の終了条件を満たすまで繰り返し実行するように制御する制御手段とを備える。従って、本発明によれば、従来例に比較して、最適な状態に向かう単調な収束を保証することができ、自由度があり、可変長のシーケ

ンスを同一のクラスで取り扱うことができ、ディジタル計算機を用いて実用的に高速処理して統計的シーケンスモデルを生成することができる統計的シーケンスモデル生成装置を提供することができる。

【０１３９】また、本発明に係る統計的言語モデル生成装置によれば、上記統計的シーケンスモデル生成装置において、上記単位は自然言語の文字であり、上記シーケンスは単語であり、上記分類手段は、文字列を複数の単語の列に分類し、上記統計的シーケンスモデルは、統計的言語モデルである。従って、本発明によれば、従来例に比較して、最適な状態に向かう単調な収束を保証することができ、自由度があり、可変長のシーケンスを同一のクラスで取り扱うことができ、ディジタル計算機を用いて実用的に高速処理して統計的言語モデルを生成することができる統計的言語モデル生成装置を提供することができる。

【０１４０】さらに、本発明に係る統計的言語モデル生成装置によれば、上記統計的シーケンスモデル生成装置において、上記単位は自然言語の単語であり、上記シーケンスはフレーズであり、上記分類手段は、単語列を複数のフレーズの列に分類し、上記統計的シーケンスモデルは、統計的言語モデルである。従って、本発明によれば、従来例に比較して、最適な状態に向かう単調な収束を保証することができ、自由度があり、可変長のシーケンスを同一のクラスで取り扱うことができ、ディジタル計算機を用いて実用的に高速処理して統計的言語モデルを生成することができる統計的言語モデル生成装置を提供することができる。

【０１４１】またさらに、本発明に係る音声認識装置によれば、入力される発声音声文の音声信号に基づいて、所定の統計的言語モデルを用いて音声認識する音声認識手段を備えた音声認識装置において、上記音声認識手段は、上記統計的言語モデル生成装置によって生成された統計的言語モデルを参照して音声認識する。従って、本発明によれば、従来例に比較して、最適な状態に向かう単調な収束を保証することができ、自由度があり、可変長のシーケンスを同一のクラスで取り扱うことができ、ディジタル計算機を用いて実用的に高速処理して統計的言語モデルを生成することができる。また、当該生成された統計的言語モデルを用いて音声認識することにより、従来例に比較して高い音声認識率で音声認識することができる。

【図面の簡単な説明】

【図１】　本発明に係る一実施形態である連続音声認識装置のブロック図である。

【図２】　図１の連続音声認識装置における単語仮説絞込部６の処理を示すタイミングチャートである。

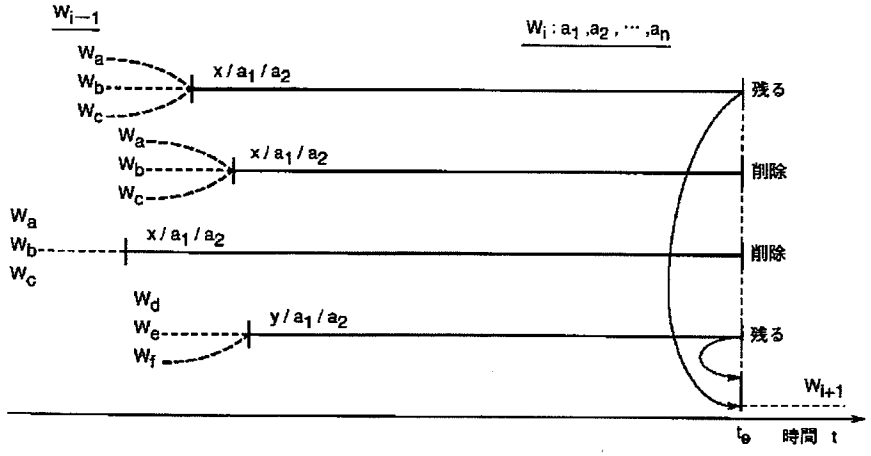
【図３】　図１の統計的言語モデル生成部２０によって実行される統計的言語モデル生成処理を示すフローチャートである。

【図４】　図３のサブルーチンであるブラウンアルゴリズムを用いた分類処理を示すフローチャートである。

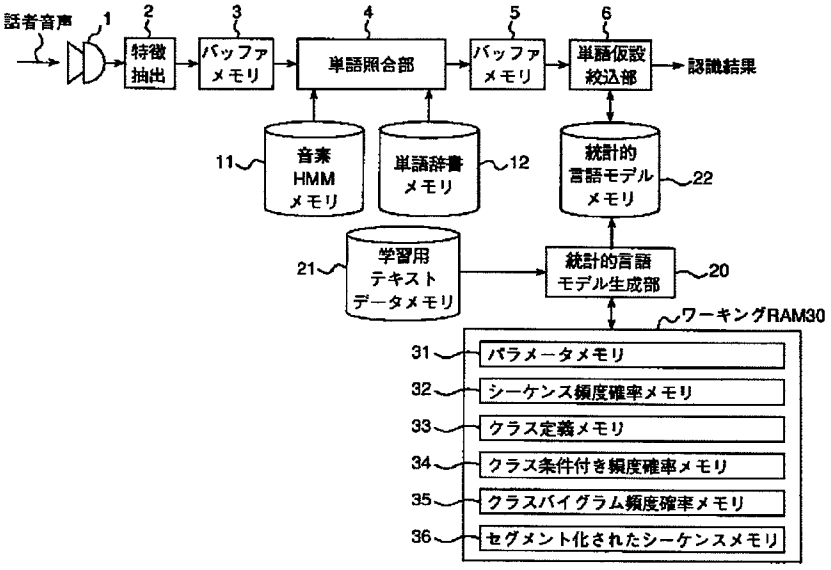
【符号の説明】

１…マイクロホン、
２…特徴抽出部、
３，５…バッファメモリ、
４…単語照合部、
６…単語仮説絞込部、
１１…音素HMMメモリ、
１２…単語辞書メモリ、
２０…統計的言語モデル生成部、
２１…学習用テキストデータメモリ、
２２…統計的言語モデルメモリ、
３０…ワーキングＲＡＭ、
３１…パラメータメモリ、
３２…シーケンス頻度確率メモリ、
３３…クラス定義メモリ、
３４…クラス条件付き頻度確率メモリ、
３５…クラスバイグラム頻度確率メモリ、
３６…セグメント化されたシーケンスメモリ。

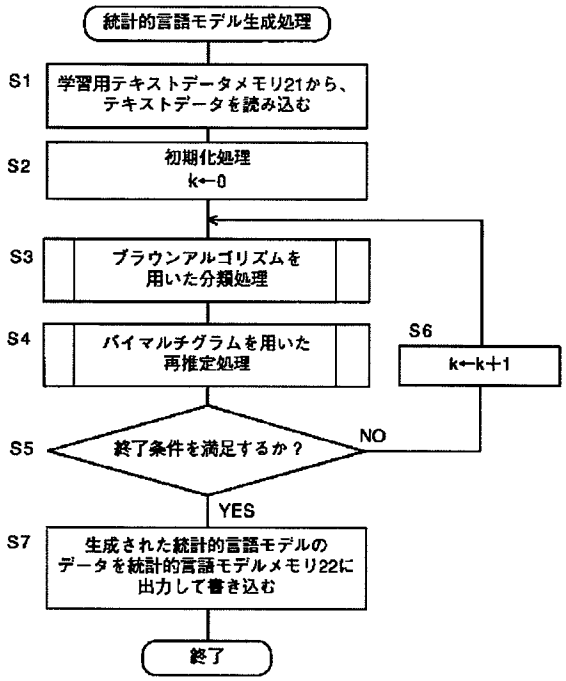
【図２】


W_{i-1}
$W_i : a_1, a_2, \cdots, a_n$
W_a
W_b
W_c
$x / a_1 / a_2$
残る
削除
削除
W_d
W_e
W_f
$y / a_1 / a_2$
残る
W_{i+1}
t_e
時間 t

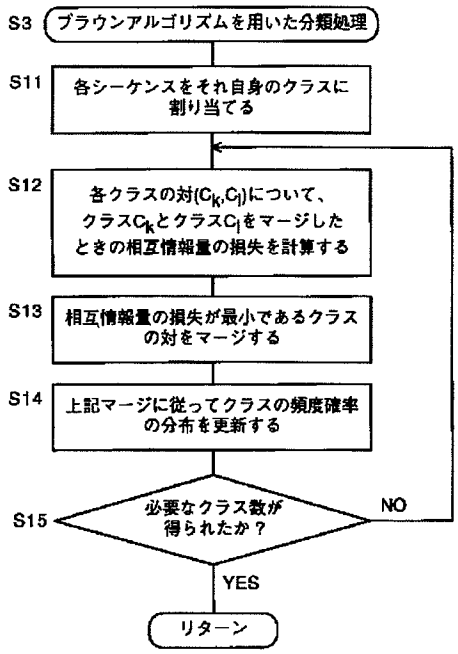
【図１】


話者音声
1
2
特徴抽出
3
バッファメモリ
4
単語照合部
5
バッファメモリ
6
単語仮設絞込部
認識結果
11
音素HMMメモリ
単語辞書メモリ
12
統計的言語モデルメモリ
22
21
学習用テキストデータメモリ
統計的言語モデル生成部
20
ワーキングRAM30
31
パラメータメモリ
32
シーケンス頻度確率メモリ
33
クラス定義メモリ
34
クラス条件付き頻度確率メモリ
35
クラスバイグラム頻度確率メモリ
36
セグメント化されたシーケンスメモリ

【図３】



【図４】



フロントページの続き

(72)発明者　中嶋　秀治
京都府相楽郡精華町大字乾谷小字三平谷５番地　株式会社エイ・ティ・アール音声翻訳通信研究所内

(56)参考文献　DELIGNE S.”LANGUAGE MODELING BY VARIABLE LENGTH SEQUENCES:THEORETICAL FORMULATION AND EVALUATION OF MULTIGRAMS”, ICASSP 1995, Vol. 1, pp169−172
Deligne S.”INFERENCE OF VARIABLE−LENGTH ACOUSTIC UNITS FOR CONTINUOUS SPEECH RECOGNITION”, ICASSP 1997, Vol. 3, pp1731−1734
Frederic B. et. al. Variable−Length Sequence Modeling:Multigrams”, IEEE Signal Processing Letters, Vol. 2, No. 6, pp111−113, JUNE 1995

(58)調査した分野(Int.Cl.[7], ＤＢ名)

G10L 3/00 - 9/20

C12N 15/00

ＪＩＣＳＴファイル（ＪＯＩＳ）